

Has the Holy See become an NGO?

The Vatican's disruption of the Cairo population conference last week is a sign that it should in future enjoy the status of just another pressure group in relation to international negotiations.

THE first week of the UN population conference at Cairo will have disappointed organizers. Despite careful negotiation, over several months, of the draft statement which it had been hoped the conference would endorse this week, the proceedings have been dominated by protests led by the Vatican over references to abortion in the document. That there might be trouble had been amply signalled in advance; the Vatican's views on abortion are well-known, and deserve respect. But its demeanour last week, and its somewhat less than successful attempt to recruit moslem countries to its cause, are more reminiscent of how pressure groups attempt to break up public meetings by obduracy rather than the intervention of a responsible member of the United Nations. The question then arises whether, by behaving as if it were a pressure group, it should be dealt with by the United Nations as if it were one — as a non-governmental organization or "NGO" in the lingo of international conferences.

The reason why the Vatican's stance at Cairo is more properly described as disruptive rather than moral is constitutional. The United Nations is, by design, a gathering of diverse states. Universality of membership is more relevant to its effectiveness than homogeneity. To the relief of the world at large, the present members include China (one of the five permanent members of the Security Council), whose strict population policy relies to some degree on abortion. But the logic of the Vatican's stance last week is that it should now campaign for the expulsion of the world's most populous state from the United Nations. (China is not the only candidate for this treatment.) That, of course, is the *reductio ad absurdum* from which even the Vatican would shrink.

That argument is not meant to undermine respect for the substance of the Vatican's complaints, which appear (from last week's reports from Cairo) to be two: all abortion is murder, while national population policies must entail the unwarrantable interference of governments in family life. The Vatican, to its credit, has been consistent on the first point. Its position is that a human embryo is a person from the moment of its conception.

The simplicity of that doctrine is admirable, but it is not everywhere accepted. Arguments about the status of pre-implantation embryos, or about the viability of embryos in the absence of a uterus, are irrelevant to the context of that statement. The fact is that, by allowing abortion to be legal under specified conditions, many governments have implicitly rejected the Vatican's position. Its redress is to persuade its adherents in the countries concerned to persuade their

governments to a different view. The Vatican has not been idle in that respect.

The second of last week's complaints is potentially more substantial, which is not to allow that the draft statement prepared for last week's conference is defective in that respect. On the contrary (see *Nature* 370, 583–584; 1994), that document sought merely to accelerate the demographic transition expected to follow increased prosperity by reinforcing the tendencies in that direction already apparent. Education, the public health of children and the "empowerment of women" were offered as chief among the reinforcing agents. That implies a bottom-up, not a top-down, strategy. The Vatican would have been better advised to plead for a statement in the final declaration of the conference that the coercion of individuals should play no part in national population policies.

The evident difficulty, for the Vatican, is that such a more constructive position would have entailed its assent to the prevalence of contraception, and of the legality of abortion, in many of the states to which the final declaration is addressed. But the sad consequence of its spurious protest last week is that the conference will have failed to flush out the opposition to the empowerment of women amply, but silently, represented at Cairo.

What will happen next? The Vatican's intervention has diverted attention from many important issues, and particularly from questions of how rapid population growth is a cause of impoverishment in the poorest of poor states. The hope now must be that something will nevertheless be done to finance population projects that developing countries welcome as instruments of development. Meanwhile, the Vatican and the United Nations separately should consider whether the Holy See should simply be registered as an NGO, a pressure group entitled to observer status at future meetings but unlicensed to be disruptive. □

The last accelerator?

If Europe cannot agree on its planned accelerator, it should internationalize the project quickly.

EUROPE seems well on the way to losing the Large Hadron Collider (LHC) planned for CERN, the cooperative high-energy physics laboratory at Geneva. Although it seemed as recently as June that the enterprise would go ahead, the



project has now degenerated to the level of intra-European squabbling about money. France and Switzerland, who both stand to benefit from the laboratory's location, have responded to a demand they they should pay more towards the final cost, but not as much as Germany and Britain had asked for. The outcome of the meeting to be held later this month will hang on whether Germany is satisfied with the extra payment (see opposite).

Opinions within the scientific community inevitably differ about the merits of this project. Particle physicists are enthusiastic, others range from the lukewarm to the sceptical. But the truth is that the LHC is the world's cheapest and quickest way of testing the validity of what is the standard model of particle physics at an energy at which the test is likely to be decisive. The cost is not so huge that the European members of CERN could not regard that as a gift to the next millennium (just over five years from now). The best hope is that the project goes ahead.

If matters continue to deteriorate, the European members of CERN should quickly bite the bullet of pride and declare that the future of the LHC must depend on recruiting other partners to full membership of CERN. Since the collapse of plans for the Superconducting Super-Collider in Texas, US (and Japanese?) membership has seemed natural. Arranging for that would take diplomacy and time and would cause delay. But that would be preferable to the abandonment of a decade's aspirations that CERN's electron-accelerator tunnel could be used for even better things. □

GATT again in limbo

The US administration has let the new tariff agreement drift too long — and dangerously.

REPORTS of the political fragility of President Bill Clinton are probably exaggerated, but probably not by much. While the US Congress prepares for the recess in which the majority of its members must fight for re-election, leaving health-care reform on the back burner for the time being, while Cuban policy is developed by daily iterations and while the invasion of Haiti has acquired (by public introspection) the status of an event that has happened, the latest alarm is that the Congress may not now pass the legislation required to give effect by the United States to the Uruguay Round of the General Agreement on Tariffs and Trade (GATT). That would be a calamity of the first order.

The immediate difficulty is that the administration has an obligation, under the legislation intended to control the federal deficit, to suggest how the reduction of federal revenues expected to follow reduced tariffs on US imports should be offset. It has not yet done so. That gives sectional opponents of provisions of the GATT agreement an opportunity to pick apart the agreement which, when ratified by its signatories, will have the status of a treaty. The danger now is that, with the November elections likely to increase the strength of the Republican opposition, opponents of the GATT agreement will be tempted to sit on their hands,

holding over the issue for the 104th Congress (which will assemble in January) and delaying the planned application of the new agreement on 1 January. If that happens, ratification by the United States will be a matter of chance.

It seems not to be appreciated in the United States what a disaster that would be. The first casualty would be everybody's prosperity. The second would be the poorest countries of the world, to which this agreement offers important routes for trading out of poverty. The third will be the willingness of other states to spend years negotiating with the United States on matters such as this — the Uruguay Round began life in 1986 — when the outcome can so easily be derailed by a balance of political forces domestic to the United States. □

Another sister journal

The impending publication of *Nature Medicine* is a welcome development of long-standing policy.

ELSEWHERE in this issue are glossy advertisements announcing the publication next January of the third of *Nature's* sister journals, *Nature Medicine*. The immediate objective, as with *Nature Genetics* (April 1992) and *Nature Structural Biology* (January 1994) is to provide *Nature* proper with a more intelligent way of responding to claims on its space from an important section of the research community, in this case those working in biomedical research. But none of this implies that *Nature* will no longer publish biomedical research, even that with evident clinical implications. The record in the past two years in genetics and structural biology is ample proof of that. But *Nature Medicine*, following the high standards of *Nature* itself, will be able to publish research at greater length (if necessary) as well as material whose importance is more obviously clinical than would be appropriate for *Nature*.

Sisterhood apart, *Nature Medicine* will make its own decisions about what to publish within the general framework of an editorial policy modelled on *Nature's*. The feasibility of that should be evident from the experience of *Nature Genetics*, which in less than two years has quarried a distinguished place in the literature. (*Nature Structural Biology* seems to be following the same path.)

Nature Medicine has other goals, notably to bridge the present gulf (such as it may be) between those whose interests are predominantly in biomedical research and those who are primarily physicians. For that reason, it will pay particular attention to the clarity of its presentation. Inevitably, there will be much about what is known by the catchphrase molecular medicine, quickly penetrating most fields of medical practice, but other techniques resting on laboratory research — magnetic resonance imaging is just one example — will also be grist to its mill. Interested readers of *this* journal will by now have guessed that the third sister-journal is unlikely to be the last and that these developments may collectively lend themselves to the electronic distribution of scientific information in interesting ways. □

German–French squabble casts pall over prospects for LHC funding

Lausanne. Hopes for an agreement later this month on the financing of Europe's next-generation particle accelerator, the Large Hadron Collider (LHC) being planned by the European Laboratory for Particle Physics (CERN), are looking decidedly forlorn.

CERN's two host countries — France and Switzerland — are still refusing to meet the demands of two other major member states, Germany and the United Kingdom, that they should pay 10 per cent of the total LHC costs in addition to their normal CERN contributions.

Last week the two countries agreed to offer a 4.3 per cent surcharge. But this level is too low to satisfy Germany and the United Kingdom. These argue that 10 per cent is the norm for countries hosting large international science facilities, for example the European Synchrotron Radiation Facility (ESRF) in France and Joint European Torus (JET) in the United Kingdom.

Burt Richter, director of the Stanford Linear Accelerator Center in California, and Hirotaka Sugawara, director of KEK, the Japanese particle physics laboratory, both warned this week that if agreement is not reached by the end of September, CERN risks losing the participation of their respective countries.

The LHC will cost CERN's 19 member

states SFr2.6 billion (US\$2.02 billion). But only SFr2.1 billion is to be paid for out of normal subscriptions. The SFr500 million deficit could be eliminated by stretching the programme over an extra two years, although the suggestion is not popular with physicists keen for large experiments to continue uninterrupted.

Christopher Llewellyn Smith, CERN's director-general, wants non-member countries to make up the difference. And after the decision by the US Congress last autumn to pull the plug on the Superconducting Super Collider, particle physicists in the United States and Japan have been turning to the LHC to fulfill their ambitions.

But Sugawara warned delegates at a meeting on large facilities in physics organized in Lausanne by the European Physical Society that CERN was developing a reputation for political infighting. Continued failure to reach agreement had prevented scientists from putting forward a convincing case to their governments, he said, and if they had to wait too long, their interest may have waned.

The LHC hit problems at CERN's June council meeting. Germany, backed by the United Kingdom, unexpectedly demanded that France and Switzerland pay a surcharge of SFr260 million in recognition of the

economic advantages the two countries obtain as host countries (CERN straddles the border between the two near Geneva). Negotiations over summer have led to a considerable increase in their initial offer of SFr25 million; but it is still well short of the German–UK demand.

The United Kingdom was prepared to compromise, and is reported to have been considering a diplomatic "redefinition" of what constitutes 10 per cent. The suggestion was that the figure could be considered as a fraction of the two countries' annual contributions of around SFr200 million over the expected 10-year duration of the programme, rather than a fraction of the whole bill. The accumulated total — SFr200 million — would represent around 7 per cent of the total costs.

Germany is said to have been much more intransigent. Its stubbornness is largely due to the financial pressures on all of its public spending since reunification. But Germany is also standing firm on the principle that as France and Switzerland reap a greater share of the benefits of CERN than other member states, they should pay extra for it.

"Principles are harder to fight than anything," says Hubert Curien, president of CERN council, who is trying to negotiate a middle ground. He believes that the host countries have greatly improved their offer, and that in any case LHC is not equivalent to ESRF and JET because it uses existing facilities (LHC will use the same 27-metre tunnel as the currently running particle accelerator LEP).

Curien suggests that, as Germany has been given a special dispensation for several years to keep its CERN contributions related to the GNP of its western states alone, rather than that of the whole country (as membership rules require), it might in turn be prepared to compromise. Both French and Swiss science are currently facing cuts in their budgets, and neither will find it easy to raise their offer.

While each side has its own troubles, the warring factions are busy blaming each other. The French say that the German demands are too high and unreasonable; the Germans accuse the French of unfairly exploiting other CERN member states by not paying for everything they get.

To limit possible further public embarrassment, Curien says that he may delay the CERN council meeting planned for 30 September until he is absolutely certain of agreement. But he insists that negotiations are not at an impasse, and remains confident of a positive decision by the end of the year. "If it does not happen, it will be very bad for physics," he says.

Allison Abbott

UK funds more science for the people

London. **The Molecule Theatre Company (right), a group which mounts performances aimed at introducing seven- to 11-year-old children to basic scientific ideas, is one of a number of organizations that will receive grants from Britain's Office of Science and Technology for furthering the public understanding of science.**

Announcing the grants last week at the annual meeting of the British Association for the Advancement of Science, David Hunt, the minister for science, added that some of the money would support a national "science week" next March, a sequel to the first such event held earlier this year.

Hunt also announced that a two-tier system, to be managed by the Royal Society through the Committee for the Public Understanding of Science, is to be established for handling awards in the field. One will be a 'seed fund' for small-scale local projects, and the other a 'development fund' for larger initiatives. The Molecule Theatre is to receive a grant of £30,000 over three years. This year's production, *The Bottle Garden*, includes a scene in which two "dolphins" (Gerard Bell and Stephen McEachran) show how the oxygen and ozone molecules are formed (see also page 192.)

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Japan backs university centres of excellence...

Tokyo. Japan's Ministry of Education, Science and Culture (MESC), is seeking several billion yen of new money in its budget request for the fiscal year 1995, submitted to the Ministry of Finance at the end of August, to stimulate the development of centres of excellence — known in Japan as COEs — within universities.

If accepted in its entirety, the new scheme will pump over ¥3 billion yen (US\$30 million) of extra money into a number of university-related institutes, and will also provide large grants of several hundred million yen each to a small number of selected teams of researchers. A relaxation of the rules on how the money can be spent will provide some much needed flexibility in the university research system.

In the past few years, every science-related government ministry and agency has sought to create COEs. MESC's scheme follows one introduced by the Science and Technology Agency last year (see *Nature* 370, 86; 1994) and the Ministry of International Trade and Industry has also introduced several measures aimed at establishing such centres. These initiatives all stem from a recognition of the need to revitalize the environment in which Japan's government-funded research takes place.

MESC's scheme consists of two parts. The ministry is asking for ¥3.42 billion to be distributed among 46 university-related institutes; these include the 14 national research institutes for joint university research, such as the Institute of Space and Astronautical Science, and the National Laboratory for High Energy Physics.

The money will go to targeted research fields in which the institutes specialize. It will be used not only for research but also to employ young research fellows on a short-term basis, to invite foreign research fellows — about one per institute — and to finance international symposiums.

Under the second part of the scheme, with a proposed budget of ¥2.69 billion for 1995, five large grants will be given to selected teams of researchers. These will be selected from researchers who are awarded 'special distinguished' or 'priority' grants by the ministry.

The special distinguished grants go to small teams of researchers headed by an individual, whereas the priority grants go to large teams of scientists spread throughout Japan. They are at present the largest and most prestigious of the ministry's grants.

The COE money will average about ¥500 million for each selected team. Unlike stand-

ard research grants, which can only be used for strictly limited purposes such as buying equipment and reagents, it can be used to employ foreign research fellows (about two per team) and young Japanese research fellows (about four per team), as well as to finance international symposiums.

Until now, grants and fellowships have been funded under separate schemes, with no all-encompassing general-purpose grants equivalent to those in the West. As a result, money can only be spent according to strict rules laid down by the ministry.

The flexibility in the new scheme is therefore likely to be welcomed. But it remains to be seen if it will lead to a more open system for employment of young post-doctoral research fellows. Nearly all research fellowships for the university system, both for foreign and Japanese researchers, are awarded through the Japan Society for Promotion of Science (JSPS), a semi-governmental organization affiliated to the ministry.

Leading Japanese researchers complain bitterly that this system stops them employing the postdoctoral research fellows they want. Details of recruitment procedures for research fellows under the new COE scheme have yet to be revealed, but JSPS will still be involved.

David Swinbanks

... but industry is sceptical of promised new tax breaks

Tokyo. Japan's Science and Technology Agency (STA) has proposed a new tax break to encourage companies to boost their research and development (R&D) budgets. But many research managers in industry are sceptical about whether such measures have any effect on the amount of money their companies spend on research in practice.

Since 1967, Japanese companies have been able to deduct from their taxable income 20 per cent of the excess of their R&D expenditure in any particular year over their previous highest annual R&D expenditure.

This system worked well for 25 years, as such expenditures increased steadily. Since 1992, however, spending on R&D has been decreasing because of the recession (see *Nature* 356, 93; 1992), and companies have not been able to benefit from the tax break.

STA's proposal was submitted to the Ministry of Finance at the end of last month, together with the agency's budget request for fiscal year 1995. Under the proposal, companies would be given the alternative option of setting 6 per cent of any increase over the previous year's expenditure against taxable income.

Although Japan is now thought to be beginning to pull out of recession, research spending by the private sector continues to remain virtually static, as confirmed by a

recent survey of companies carried out by the *Nikkei Shimbun*, a leading financial newspaper (see table).

The STA hopes the 6 per cent tax break — which, if accepted by the Finance Ministry, will become effective in fiscal year 1995 — will encourage companies to spend more on research. But research managers in Japan's private companies, while welcoming the proposal, are sceptical.

Ranking		Amount ¥ billions
1	NTT	300 (+3.6%)
2	NEC	280 (0.0%)
3	Toshiba	275 (-0.1%)
4	Fujitsu	270 (-1.5%)
5	Mitsubishi Electric	165 (-0.6%)
6	Mitsubishi Heavy Industries	117 (+1.5%)
7	Sharp	110 (+1.6%)
8	Sanyo Electric	81 (-0.2%)
9	Fuji Photo Film	73 (+2.0%)
10	Mazda Motor	72 (+3.3%)

Source: Nikkei Shimbun
Figures in parentheses represent % change 1993-1994

In 1985, the government introduced another tax incentive for research, offering a 7 per cent tax deduction on money spent on equipment and facilities for basic research. The move coincided with the creation of a number of basic research institutes by the private sector in Japan, and some have attributed this in part to the tax incentive.

But many in industry and elsewhere argue that the 1985 tax break simply followed an existing trend by companies to set up basic research institutes during the bubble economy of the mid- to late-1980s, when these companies were flush with money, and cannot be seen as the cause of the trend.

The STA claims that its proposed new tax break has been put forward in response to prompting by industry. An official of the Ministry of International Trade and Industry argues that such tax incentives have been an important factor in encouraging the large expenditures on R&D by the private sector.

But the manager of research planning at one company says that industry only made requests to the STA after it had been asked to do so by the agency. He claims that the STA proposal is merely an attempt by government officials to appear to be trying to do something about research spending in industry — and that such moves have little impact on decision making in industry. **D. S.**

NASA buys into the action on Keck telescope

Washington. The US National Aeronautics and Space Administration (NASA) is close to signing an agreement with the California consortium that operates the Keck telescope, the largest optical instrument in the world, under which the space agency will pay one-sixth of the Hawaiian observatory's construction costs in exchange for an equal proportion of viewing time.

NASA will use the two 10-metre Keck telescopes, the second of which is scheduled to be completed in 1996, for a long-term programme to search for planets around other stars. One-third of the viewing time on the Keck II telescope will be dedicated to that search.

The deal, due to be finalized within the next several weeks, calls for NASA to pay \$6.8 million this year, with equal amounts paid out over each of the next five years, to a total of about \$40 million. NASA's investment will be the only government money received by the observatory, which is financed largely through a private grant from the W. M. Keck Foundation.

The first Keck telescope began regular operations last year. The identical Keck II is being built 85 metres away on the peak of Mauna Kea, the Hawaiian volcano that offers astronomers one of the best viewing sites in the world. The telescopes are designed to be used both independently and in combination as an optical interferometer, which will greatly improve their joint resolving power.

According to Edward Stone, director of NASA's Jet Propulsion Laboratory and chairman of the Californian research consortium that manages Keck, it is partly the interferometry capability that makes NASA's involvement so attractive. The space agency is spending its own money to develop new techniques for optical interferometry in a test-bed project currently being carried out at Mount Palomar in California. Those techniques, to be used in the Astronomical Studies of Extrasolar Planetary Systems (ASEPS) programme, will also benefit other astronomers who are planning to use Keck.

The eventual goal of the ASEPS project is to detect Earth-sized planets around other stars, if they exist. But this is likely to require space-based instruments. The first phase of the programme, using the ground-based observations from Keck, will concentrate on indirect detection of large planets around stars that are relatively near.

The presence of planets with the mass of Uranus or Neptune can be inferred from spectral or astrometric study of a star, whose motion will be affected by the gravitational tug of a nearby massive object. ASEPS will also study protoplanetary disks — solar systems in the process of formation — which have already been observed around several

stars. Keck may also be able directly to image Jupiter-size planets around nearby stars.

An outside advisory group was due to meet in Washington this week to advise NASA on how the ASEPS programme should be run. The current plan, according to Jurgen Rahe of the agency's solar system exploration division, is to establish a core team that makes on-going observations, while funding individual investigators to conduct specific research projects in extrasolar planet detection.

Bavaria bids for key role in biotech

Munich. The southern German state of Bavaria, already known as Germany's Silicon Valley, could also become the biotechnology capital of Central Europe, despite the country's traditional hostility to such techniques, according to the state's economics minister, Otto Wiesheu.

Last week, Wiesheu announced the creation of a venture capital fund to help establish new biotechnology companies. The fund forms part of a campaign launched last July to promote new technologies in the state.

Bavaria's ministry of economics is supplying DM150 million (US\$97.35 million) to found a venture capital company for key technologies, and it wants German banks to match this amount. DM50 million will be earmarked for biotechnology.

The new money comes on top of a DM28-million fund, financed almost entirely by the privatization of the Bavarian utilities

Until recently, the ASEPS programme was known as TOPS (Toward Other Planetary Systems). But NASA managers, worried that TOPS had become tainted when Congress cut the Search for Terrestrial Intelligence (SETI) programme from the agency's budget last year, decided to change the name. They say they are confident that the astronomical search for other planets, unlike SETI, should be safe from political attack — or at least as safe as any other project in Washington. **Tony Reichhardt**

donated by the Max Planck Society, Germany's principal basic research organization.

Martinsried is already a major centre for biological research, being home to two large Max Planck institutes, several university departments, and a large teaching hospital. Munich's Ludwig Maximilians University will be moving its entire chemistry and pharmacy faculties there in 1999.

The Centre for Biotechnology will rent out laboratory and office space — initially at subsidized rates — to newly established biotechnology companies or research laboratories.

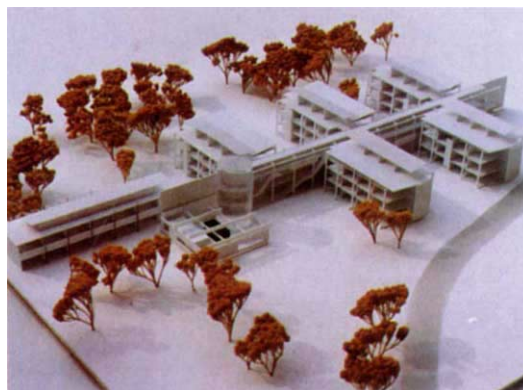
Building will start at the beginning of next year, and there will be sufficient space for at least 10 companies to move in when the first phase is completed by the end of the year. The second phase will be completed in 1996, offering space for a further 30 companies. The centre will provide major items of expensive equipment, such as sequencers and centrifuges, which companies may share without cost.

At present Germany's biotechnology sector consists of only 10 small firms. But recent amendments to Germany's tough laws on genetic engineering (see *Nature* 367, 210; 1994), combined with campaigns by science organizations, have improved the prospects for biotechnology.

"The anti-biotechnology movement in Germany has softened," says Axel Ullrich, head of molecular biology at the Max Planck Institute for Biochemistry at Martinsried, who was involved three years ago in setting up Sugan, a small biotechnology company jointly founded by the Max Planck Society and New York University.

"At the time we could no way think of setting it up in Germany," he says. Instead, Sugan based itself in California, but Ullrich says it is now considering installing a spin-off company in the new Martinsried centre.

Alison Abbott



Architects' model of new Martinsried centre.

company Bayernwerk-AG, which will be used to build a large biotechnology centre in the state. DM5 million of the building fund will come from the state budget.

The programme to build up biotechnology activities in Bavaria has moved rapidly since Edmund Stoiber, the state's prime minister, announced the plan to parliament last summer. By autumn, a site in Martinsried near Munich had been chosen on which to build a 5,600-square metre 'Centre for Biotechnology', on land

Brussels seeks panel's help to link European research

Brussels. The European Commission last week launched an experiment in power-sharing with the scientific community, with the inauguration in Brussels of the European Science and Technology Assembly (ESTA), a high level advisory body of 100 European scientists (see *Nature* 368, 385; 1994).

The main task of ESTA is to improve the quality and choice of European Union (EU) research programmes, says Antonio Ruberti, the research commissioner. The commission's other hope, he says, is that the existence of the assembly — which includes the heads of several national research agencies and six Nobel prizewinners — will help generate better coordination between national and EU research programmes.

The commission's dilemma here is that it is mandated to improve coordination by the 1991 Maastricht agreement, but only accounts for around four per cent of total spending by EU member states on research. Moreover, national research bodies (and their governments) are loathe to concede more power to Brussels.

Indeed, the creation of the assembly comes at a time when the commission badly needs allies. The 12 EU governments, represented in Brussels by the Council of Ministers, see an opportunity to claw back power from the commission, now that its bruising president, Jacques Delors, has been replaced by the less confrontational Jacques Santer, formerly prime minister of Luxembourg.

EU governments also hope to strip yet more powers from the commission at a review conference of the Maastricht Treaty scheduled for 1996, in line with a general view that political control in Europe should remain as decentralized as possible.

Research is a relatively small part of the commission's work, but one of its most controversial. Some argue, for example, that the EU should switch funding away from 'pre-competitive' research to focused social and commercial goals, as part of a 'shock therapy' industrial policy.

What direction EU research takes will depend heavily on the outcome of ongoing negotiations on the form of the incoming commission. Some advocates of industrial policy favour merging the directorate for industry (DGIII) with the research activities of research and education (DGXII), particularly as the use of research funds as explicit subsidies to industry has been made easier by the recent GATT agreement.

But another prospect would be the splitting of research and education into independent directorates; if Sweden, Finland, Norway and Austria join the EU as expected, the commission will need to slice its salami more thinly to provide jobs for their

commissioners.

Much will also depend on who replaces Ruberti as research commissioner at the end of the year. Even if a merger between DGIII and DGXII were excluded, for example, the possible appointment of a commissioner from a country such as France, which is keen on industrial policy, could amount to much the same.

Whatever direction the commission takes, it hopes that allowing the assembly to comment on all proposed EU programmes will

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Ruberti: aiming for better coordination

strengthen its hand in subsequent negotiations with the Council of Ministers, as such proposals will now appear to carry the endorsement of the research community.

Some government officials in member states say, however, that they will not be moved by the "rubber stamp" of a body seen as the commission's "poodle". Whatever happens, the council will increasingly dominate the commission, says one official at the French science ministry. When EU governments meet in 1996, for example, they will probably try to reduce the powers of the commission's main ally, the European Parliament, by trying to abolish its current veto on the research budget.

Another tactic of the commission may be to use the assembly to increase its influence over national funding agencies and European-level bodies, such as the European Space Agency as it will be more difficult for them to criticize decisions they have been involved in.

National agencies are seen by some in Brussels as a potential threat to the commission's role in overseeing and shaping European research, giving as examples pressures to devolve management of research away from Brussels and the recent decision of several national agencies to create their own transnational 'associated European laboratories' without the help of the commission.

Such a tactic could backfire, however, says one commission official, who talks of a 'Frankenstein scenario' where the commission, having created an assembly of powerful individuals close to itself, might subsequently find its power base being eroded from within.

Declan Butler

Ireland agrees to cover tax costs on Wellcome grants

Munich. The Wellcome Trust, Britain's largest medical research charity, has withdrawn a threat to discontinue its support for research in Ireland after reaching agreement with the Irish government in a dispute over the payment of value added tax (VAT) on research equipment.

After prolonged negotiations, the Irish government has now agreed, through its Health Research Board, that it will in future cover the costs of VAT on Wellcome grants in a venture it describes as "co-funding".

The trust had challenged its obligation to pay VAT at a rate of 21 per cent on all such equipment bought for Irish grant recipients. It had argued that all of its money should go to research without being taxed in this way.

The problem had surfaced as the result of three major programme grants awarded by the Wellcome Trust to Irish scientists since 1990. Previously, grants from the trust amounted to only around £100,000 (US\$153,300) a year; but as a result of the new awards Wellcome's funding of research in Ireland has risen to £5.4 million over the past four years, and it has faced a £400,000 bill for VAT.

In response, the government had argued that it had little room to manoeuvre within taxation rules introduced by the then European Economic Community in 1989, forbidding members from creating new exemptions to VAT liability. (In the United Kingdom, such equipment was already exempt before the new rules were introduced.)

Earlier this year, after the trust had threatened to withdraw all funding if the tax demand was sustained, the government announced that items of equipment for medical research bought by charities at a cost of more than £20,000 would be liable for a VAT refund.

But this did not satisfy the trust, as few single items of equipment cost more than £20,000. For example, Brian Harvey, professor of physiology at the University of Cork, says that he spent £290,000 on equipment this year out of a grant from the trust — but that the most expensive single item only cost £19,000.

● Seamus Brennan, Irish minister for Commerce and Technology, is due to make an interim statement on 19 September about the work of the Science, Technology and Innovation Council, a body of 18 scientists, industrialists and civil servants, which was set up in February to review Irish science policy. Brennan's statement is expected to include a promise of more funding for research; the government was heavily criticized last year for failing to provide any new money for non-medical basic research.

Allison Abbott

US scientists seek support for nuclear sub

Houston. Wanted: one decommissioned US Navy nuclear submarine for use as a dedicated scientific research vessel. This may be the ultimate in dual-use concepts now being explored in the wake of the Cold War. But it will be top of the agenda at a scientific workshop being held next week in Washington to gauge the level of interest in scientific and government circles for what has been coined the 'white submarine' concept.

Meeting organizers hope that the workshop will also allow them to pin down the scientific rationale for using this type of vessel for scientific research primarily in — though not limited to — Arctic waters.

The scientists backing the idea emphasize that they are not advocating the use of a nuclear submarine for research that could be done using surface ships or other means. Nuclear submarines come into their own when there are "hostile conditions at the surface, be it ice or storms", says Lloyd Keigwin, a senior scientist at the Woods Hole Oceanographic Institution in Massachusetts, and one of the organizers of the two-day workshop.

A nuclear submarine, he says, would enable researchers to map the sea-floor in the Arctic or Southern Oceans, to study global climate change, to improve our understanding of basic ocean circulation, or to track environmental pollution. "Understanding upper ocean physics and sea floor processes in the Arctic and Antarctic are just two very fundamental things that we'll never be able to do any other way," says Keigwin.

Next week's meeting will bring together researchers from a broad range of scientific disciplines, including chemists, geologists, geophysicists, marine biologists, physical oceanographers, 'acousticians' and specialists in submarine technology, as well as government officials.

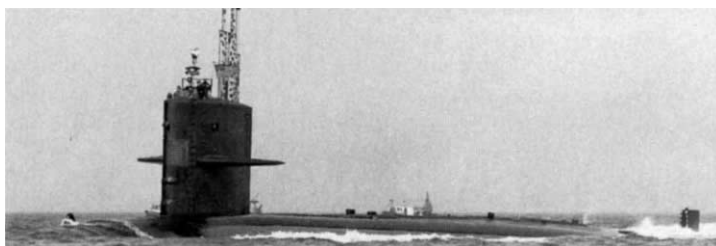
Keigwin hopes that it will produce a set of research priorities (and geographical locations) that will form the basis of a 'white paper' to follow on from the SOONS (*Scientific Opportunities Offered by a Nuclear Submarine*) report, published in 1992 by the University National Oceanographic Laboratory System.

What makes the matter more urgent is that the entire fleet of nuclear submarines considered to be most appropriate — the *Sturgeon* class, which first came into service in the 1960s — is due to be decommissioned by the US Navy over the next decade or so. Special design features of these submarines, in particular toughened sails and diving planes that can be rotated to the vertical

position, enable them to break through the Arctic icepack to surface should the need to do so arise.

"We have this one-time window of opportunity, as these boats go out of service, to use them for a little while before their reactors are worn out," says Garrett Brass, executive director of the US Arctic Research Commission. "That's why we're anxious to strike now while the iron is as hot as it will ever be," he says.

Cost estimates for a white submarine



A *Sturgeon*-class submarine: new role in a post Cold War world?

vary enormously. After reinforcing scientific support, a detailed cost study would probably be the next task, says Keigwin. Estimates for overhauling a nuclear submarine range from \$50 million to three times that amount; running costs could range between \$8 million and \$12 million a year.

The question of who would foot the bill also remains open. Costs of that order of magnitude make it "impossible for any one agency to fund", says Gary Hill, chief of the Office of Energy and Marine Geology at the US Geological Survey (USGS), which has responsibility for monitoring and mapping the sea-floor and sea-bed. But Hill, who argues that the project would support USGS's mission, thinks that one of the keys is going to be the Navy's reaction.

The Navy itself has yet to be convinced. David Albritton, a spokesman for the serv-

ice, says that the white submarine proposal is "currently being evaluated" and that it would therefore be "premature" for the Navy to comment on the scientists' suggestions. He declined to say who, if anyone, from the Navy would be attending next week's workshop.

Although there is no precedent in the United States for a white submarine, Russia has on several occasions offered to convert a nuclear submarine from its own fleet into a research vessel for use in the west. Indeed, Keigwin was a member of a US delegation that went on a fact-finding mission in 1992 to the Russian Submarine Design Bureau in St Petersburg.

This, and other similar initiatives, proved to be a political hot potato and came to nothing. But there are indications that the US Navy might slowly be warming to the idea of making its military hardware more accessible to the civilian science community at large. Last summer saw the first unclassified science mission aboard a fully operational nuclear attack submarine, the *Pargo*, which sailed from Groton, Connecticut, to the Arctic with five civilian scientists on board.

The *Pargo* mission was widely seen as a success, and by some as the litmus test for future missions. The Navy has committed itself to five more dedicated science missions, each lasting 45–60 days, over the next five years. It recently signed a memorandum of understanding along with several government agencies that include the National Oceanic and Atmospheric Administration, the National Science Foundation and USGS. The next mission to the Arctic is set for March.

Diane Gershon

Pearson resigns as head of Darwin Corp.

San Francisco. Only months after three high-powered investors joined its board, the president of Darwin Molecular Corporation has resigned over "strategic differences". Mark Pearson, president and chief executive of the small Seattle-based gene sequencing company — and previously an executive director at DuPont Merck Pharmaceutical Co. — blamed conflicts with the board for his resignation, but declined to elaborate.

In May, William Gates III and Paul Allen, the co-founders of Microsoft Corp., invested \$10 million in Darwin. They joined the board along with George Rathmann, chairman and chief executive of Icos Corp. in Seattle, and chairman emeritus of Amgen Inc. (see *Nature* 369, 88; 1994).

The company says that Pearson's departure will not lead to any change in its scientific focus or technical strategy. Darwin is using a three-pronged approach to drug discovery, involving gene sequencing, the computer-aided analysis of sequencing information, and 'test-tube evolution', which subjects generations of targeted molecules to a test of survival.

Pearson said Darwin is preparing to announce a significant additional cash infusion by a large group of investors. David Galas, who came to the company from the Human Genome Project at the US Department of Energy, will act as interim chief executive until a replacement for Pearson is appointed.

Sally Lehrman

Anti-AIDS strategy targets immune system

San Francisco. The National Institute of Allergy and Infectious Diseases (NIAID), frustrated with traditional, small-molecule, drug-based approaches to AIDS, has launched a programme to bring research institutions together for work on novel, immune-based strategies.

NIAID said it would dedicate \$25 million over the next four years to the Strategic Program for Innovative Research on AIDS Treatment (SPIRAT) for approaches such as restoration of the immune system, gene therapy and DNA-based therapeutic vaccines.

"This effort is crucial because currently available anti-HIV drugs only partially and temporarily suppress replication of the virus," says Anthony Fauci, director of NIAID.

AIDS treatment activists said SPIRAT, although a small part of NIAID's \$558 million AIDS budget, represented a major shift in focus that has been long overdue for the US agency. Even conservative AIDS researchers are beginning to shift their focus to immunological strategies as it becomes clear that antivirals alone will not be effective, said Mark Harrington of the Treatment Action Group in New York.

Activists lamented, however, that the funding appears to come at the expense of another programme, the National Cooperative Drug Discovery Groups for the Treatment of HIV (NCDDG). That effort, which fostered industry collaborations with academia, is being scaled down.

David Ho, director of the Aaron Diamond Research Center in New York and a member of the National Task Force on AIDS Drug Development, said any shutdown of NCDDG would leave a void. Pharmaceutical companies with the tools to continue the pursuit of small-molecule drugs need incentives to remain in AIDS research, he said.

Nava Sarver, an acting branch chief in the division of AIDS at NIAID, said that set-aside funds for NCDDG had been discontinued. But researchers can still apply for investigator-initiated, unsolicited support under the NCDDG framework. "We are very, very eager to continue the programme," Sarver said.

NIAID is encouraging industry participation in the new grants, and four of the six university-based researchers with initial SPIRAT grants have corporate collaborators. The teams are to receive about \$1 million a year for four years, with early studies in humans to start by the third year.

Harrington and other activists said the new programme's inter-institutional approach would help bridge a gap in AIDS research between the laboratory and the clinic by encouraging greater interaction between clinicians and those working at an

early stage in the research.

"There's been a brick wall between the two," said Martin Delaney, an activist with Project Inform in San Francisco and a member of the AIDS Research Advisory Committee of NIAID. All too often treatment approaches die because of poor communication between basic scientists and clinicians, he said.

The projects include several techniques to bolster the immune system response by manipulating T cells, adding cytokines or infusing cells from non-infected sibling donors. Gene therapy strategies would concentrate on a defective gene from HIV for use in children, and an antiviral gene to produce ribozymes that would inactivate part of HIV. A vaccine-based approach would inject non-infectious HIV genes into the muscle tissue of HIV-infected people.

Sally Lehrmann

US blocks return of reactor fuel

Munich. South Carolina has placed an injunction on the US Department of Energy to stop it bringing highly enriched uranium (HEU) waste from research reactors in Europe for storage at the Savannah River Site in South Carolina.

Waste from four research reactors — in Holland, Sweden, Norway and Austria — was dispatched by ship last week. But a day later, South Carolina's governor, Carroll Campbell, filed a suit against the Department of Energy, claiming that the European countries should store the waste on their own territory, and that a proper evaluation of the impact of the waste on the state's environment had not been carried out.

The move comes as a major blow to the

US anti-nuclear proliferation agency, the Nuclear Control Institute (NCI), which has been actively engaged in persuading the four research reactors to convert to the use of low enriched uranium (LEU) (see *Nature* 369, 89; 1994).

Part of the agreement was that the United States would handle all radioactive waste from the reactors immediately, as each was close to the limits of their licences for storing waste at their own sites. The Department of Energy's approval of immediate dispatch of the waste was received in June.

Marcel de Bruin, director of one of the research reactors at Delft in the Netherlands, says that if the department is unable to mount a successful challenge to the injunction, he will have to consider sending his waste to the United Kingdom for reprocessing — a move which would contravene current policy in both the United States and Holland.

De Bruin said he would also have to reconsider the agreed expensive conversion to LEU. "If these research reactors start to take such steps", says a spokesman for the Nuclear Control Institute, "the whole policy on controlling nuclear proliferation could collapse".

The National Resources Defense Council, an environmental group based in Washington, has joined the Nuclear Control Institute and the research reactors in challenging the injunction. All three have criticized the timing of Campbell's injunction, pointing out that it was issued after the two ships bearing the highly radioactive cargo had set sail, and left the cargo potentially vulnerable to accident or terrorist attacks.

Alison Abbott

CNRS puts lab equipment orders in suspense

Paris. Sun-tanned and relaxed after the traditional long French summer break, laboratory heads returning to the Centre National de la Recherche Scientifique (CNRS) institutes throughout France may have turned pale on seeing the message feeding through their fax machines from the newly appointed director-general, Guy Aubert: "suspend all orders paid out of CNRS funds".

Aubert, barely in the job two months, has moved swiftly to avert an impending financial crisis at CNRS. In a bid to obtain an accurate picture of CNRS finances, he has also asked laboratory heads to supply him with a list of all orders already passed, and estimates for spending until the end of the year. Only then will he let staff know how much they can spend.

Researchers seem to have accepted the action as inevitable — and even overdue. The financial problem has arisen because CNRS laboratories spend money according to promises made by the government in the annual budget, whereas actual cash payments often lag behind.

Indeed, while Aubert's action is an emergency measure, he is keen to end what he describes as obsolete financial management at CNRS. He plans to create a new system, for example reducing the number of staff authorized to make equipment purchases and increasing the accountability of laboratory heads. Last year, many suppliers went unpaid until laboratories received their 1994 budget; Aubert believes such a situation is bad for CNRS's image.

Declan Butler

UK urged to open up science advisory panels

Loughborough, UK. Pressure is building on the British government to increase the public's acceptance of science-related activities — for example, those relating to the applications of biotechnology and genetic engineering — by opening up its science advisory bodies to greater public scrutiny.

The pressure has highlighted differing, and occasionally conflicting, goals between those currently promoting the increased 'public understanding of science'.

Last week, David Hunt, in his first major speech as the new minister for science, told the annual meeting of the British Association for the Advancement of Science (BA) at the University of Loughborough that the government had decided to provide substantial extra support for various activities aimed at achieving this goal (see page 187).

Justifying these moves, Hunt said it was essential to stimulate a "proper dialogue" between government, industry and the scientific community — and between specialists and citizens. "We need to generate an excitement about science and the scientific world for future generations," he said.

But Hunt reacted coolly to a suggestion by the *Financial Times* newspaper that what was also needed was increased openness of government advisory committees (the US National Institutes of Health's Recombinant DNA Advisory Committee, whose meetings are generally held in public, was quoted as an example to emulate).

The newspaper's proposals have some strong supporters in private industry. Con-

cern is growing among biotechnology companies, for example, that the traditional secrecy which still surrounds much of the government's advisory machinery may itself be helping to generate a distrust towards their activities.

"If the legal process for getting clearance was more open, we believe that a lot of the current opposition would go away," says Louis da Gama, executive director of the BioIndustry Association. He points, for example, to the recent controversy surrounding the experimental release of a modified caterpillar virus carrying a scorpion toxin gene in a field in Oxfordshire (see *Nature* 369, 348; 1994).

A similar message seems likely to emerge from a 'consensus' exercise currently being carried out by the Science Museum on behalf of the Biotechnology and Biological Sciences Research Council, in which a lay panel has been set up to discuss in detail potential public concerns about plant biotechnology (see *Nature* 369, 435; 1994).

The conclusions of the panel, which has already met for a weekend of informal discussions, will not be finalized until a two-day public conference at the beginning of November. But according to John Durant, assistant director of the Science Museum, a key factor already emerging is the need to ensure the integrity and credibility of the regulatory process.

So far, however, the government appears to be relatively unmoved by demands that this would be enhanced by greater open-



ness. Asked at a press conference for his reactions to the *Financial Times* editorial, Hunt said that he wanted to see the whole process as open as possible — but that "on particular topics there is a need to call together experts in a private way".

The debate seems destined to intensify as government-funded efforts to increase the public understanding of science come under increasing scrutiny, inevitably including the motivations of those involved.

Jill Nelson, for example, head of science promotion at the Royal Society, told a separate meeting that scientific learned societies and professional bodies must think carefully about the relationship between their desire to promote science as a public duty, and that to push their corporate identity.

A survey of members of STEMRA — the recently formed Science, Technology, Engineering, Medicine Public Relations Association — showed that most professional societies combine their public understanding of science activities with their corporate press and public relations, with the same individuals often taking on both responsibilities because of limited resources. "When the society confuses the two objectives, it confuses the messenger and the message," said Nelson.

There was also a warning of the hazards facing social scientists attempting to increase their understanding of the way that scientists operate in a sharp exchange between Lewis Wolpert, the chairman of the Committee on the Public Understanding of Science, and Harry Collins, director of the Science Studies Centre at the University of Bath.

Having strongly criticized sociologists for the relativism of their theorizing, Wolpert at one point suggested that their main motivation was "envy" of their natural science colleagues, and that they suffered from a "massive inferiority complex". In return, Collins claimed that Wolpert was "engaging in a pantomime".

David Dickson and Maggie Verrall

Demand grows for 'positive' gene therapy

Upbeat media coverage of the first UK gene therapy trials for cystic fibrosis may have contributed to an almost threefold increase over the past year in public support for the use of genetic techniques to enhance desirable traits in children — a trend which is already generating some concern among British researchers.

A recent Gallup poll, commissioned by the psychology and genetics research group at Guy's Hospital in London and funded by The Wellcome Trust, showed that three times more people would now consider changing the appearance and behaviour of their children through the implantation or alteration of selected genes than admitted so a year ago in a similar poll conducted for *The Daily Telegraph*.

The overall figures remain relatively small. But this shift is already raising questions over how far the views of the general public should be taken into account in deciding policy guidelines on the application of gene therapy. "Do you want to incorporate public opinion where the vast majority might be eugenic?" Theresa Marteau,

director of the group at Guy's Hospital, asked last week's meeting of the British Association for the Advancement of Science.

The numbers of positive responses to questions about such uses of gene therapy were still small, with no more than 20 per cent in any one case saying that they would use the technology if it were possible. But it is the rapid rise in acceptability over the past year that is potentially more significant.

For example, the percentage of respondents saying that they would use genetic methods to affect aggression rose from 5 per cent in 1993 to 18 per cent in 1994. The corresponding increase for alcoholism was from 5 to 18 per cent, for homosexuality from 4 to 10 per cent, and for 'good looks' from 2 to 5 per cent.

Marteau suggested that the large amount of coverage about the gene therapy trials for cystic fibrosis may explain why the public has become "more positive" about the uses of gene therapy, even though they remain cautious.

Maggie Verrall

World government the answer?

SIR — Following your identification of the need to strengthen international political and economic institutions (*Nature* 370, 83; 1994), I wish to draw your attention to a recent proposal for increasing the democratic accountability, and therefore the political legitimacy, of the United Nations (UN). The proposal, put forward originally by the World Federalist Movement, calls for the establishment of a UN Parliamentary Assembly, involving people appointed from national parliaments, as was originally the case for the European Parliament (and in its early form, the US Senate), that would, like them, eventually evolve into a directly elected body.

As at present constituted, the UN represents the interests of governments, especially those of the five permanent members of the Security Council. But its activities increasingly impinge directly on the lives of many millions of world citizens who are not represented in its decision-making process. This occurs, for example, through the imposition of punitive economic sanctions (Haiti, Iraq, Serbia), UN-sponsored military action (the Gulf War), peace-keeping operations (Bosnia, Cambodia) and attempted disaster relief (Somalia, Rwanda). In these and other ways, the UN is already trying to act as a world government and, as highlighted by your leading article, there is indeed an urgent need for this role to be developed further. However, it is clear that the lack of democratic accountability is impeding the UN's ability to fulfil this role.

Ever since John Locke wrote his *Two Treatises of Government* (1689), it has been generally accepted (at least in the West) that if the lives and liberties of individuals are affected by the actions of government, they must be represented in that government. As Locke warned: "When anyone . . . shall take upon them to make laws, whom the People have not appointed so to do, they make laws without authority, which the People are not therefore bound to obey" (*Second Treatise of Government*, para 212).

World government will be no different from any other kind of government in this respect, and, unless a way is found of giving the world's citizens a voice in its activities, the UN will continue to fail to meet many of its most important objectives owing, at least in part, to a perceived lack of political legitimacy.

The establishment of a UN parliamentary assembly, even if it initially had only a consultative role, would go a long way towards correcting the present democratic deficit in the organization of the UN, and therefore enhance its authority. In the longer term, such an institution might form the foundation for a genuinely democratic federal world government.

Finally, as a matter of practicality, it should be noted that a parliamentary assembly could probably be established in the short term as a "subsidiary organ" of the General Assembly (under Article 22 of the UN Charter), thereby avoiding the great difficulty of having to renegotiate the charter itself. Further information can be obtained from the World Federalist Movement, UN Office, 777 UN Plaza, New York, New York 10017, USA.

I. A. Crawford

1 Cochrane Road,
London SW19 3QP, UK

Research squeeze

SIR — After thirty years in medical academia, I'm not baffled by the 54 per cent drop in National Institutes of Health (NIH) grants for young scientists over the past eight years (*Nature* 370, 87; 1994). When asked why he robbed banks, Willie Sutton replied: "Because that's where the money is". Now known as Sutton's law, it is the most parsimonious explanation for the NIH fall-out.

The name of the game may be research but the real bottom line is money; money to support the investigator directly and, thereby, indirectly to support the investigator's institution. In the 1980s, NIH monies dried up relative to demand. For academic physicians, particularly the most junior (and least tenured), the way of soft money became a form of Russian roulette few could chance. Academic clinical practice generated hard money offering *de facto* tenure as secure, if not as prestigious, as the institutional tenure possible through research (grants).

So it is and it will get worse. Over the past few years, the economic pressures for managed health care have come to dominate most academic medical institutions in the United States. These pressures barely tolerate the costs of teaching, let alone research. Given this reality, the possibility of a research career for young academic physicians must continue to shrink to an option for a few, mostly in select institutions. The economics of managed health care will similarly affect career opportunities for PhD investigators in academic medical institutions. As soft money to support PhD research becomes scarcer, it must be replaced by hard money generated through institutional clinical practice plans. Such largesse is not the mentality of managed health care.

Mark W. Steele

University of Pittsburgh,
School of Medicine,
3705 Fifth Avenue at Desoto Street,
Pittsburgh,
Pennsylvania 15213-3417, USA

Life class

SIR — Recent correspondence on the report of the "Pontifical Academy for Life" (*Nature* 368, 90; 1994) is reminiscent of the 13-year-old scientific controversy aroused by the US Senate hearings on "Human Life" (*Science* 212, 648; 1981). In this respect, the letter by J. A. Botella (*Nature* 370, 172; 1994) shows the lasting inconsistency of the Vaticanist arguments against contraception and abortion: (human) life, it is said, begins at the moment of conception, when the genetic material of two germ cells contributes to create a new individual, different from its parents. Then, as reiterated by Professor Jerome Lejeune's followers, the human nature of the human being from conception is not a metaphysical contention, but plain experimental fact. The fallacy of the assertion lies in embryological as well as in etymological and semantic grounds.

As stated by B. G. Boving (*Science* 213, 154; 1981), a fertilized egg is a single cell (zygote) prone to a series of developmentally regulated divisions bringing out two major prime components: the embryo-blast (mostly bound to be a fetus, or two identical twins) and the trophoblast (which develops into membranes, placenta, and umbilical cord). All these cells and components are alive, human and have the same genetic mixture, but they can hardly be seen as a person or an individual human being (non divisible single entity). Hence the etymological fallacy, which also involves a metonymical substitution of human cells for human person. Given the weakness of the scientific arguments, it is a surprise that Botella has to invoke some *mysterious* events to justify the sudden appearance of a person out of a fertilized egg. I suggest that the mystery of humanhood be rated first in the scientific agenda of the "Academy" for Life.

Angel Pestaña

Instituto de Investigaciones Biomédicas,
CSIC, C/Arturo Duperier 4,
280029 Madrid, Spain

Second sex?

SIR — Apparently, no women have contributed enough over the 125-year history of scientific excellence at *Nature* to warrant a female speaker at your celebration. The lack of female speakers is emphasized by the title of your symposium, "Our place in nature — the evolution of our world." You must be talking about a male-dominated world that evidently has *not* evolved over the last 125 years.

Georgiana May

Department of Plant Biology,
University of Minnesota,
1445 Gortner Avenue,
St Paul, Minnesota 55108-1095, USA

Russian science in the balance

Vladimir Pokrovsky

Everybody knows that the situation for scientific researchers in the former Soviet Union is dire. But, in Russia at least, there is some reason for optimism.

HAVING depended for decades exclusively on the government for its resources, basic science in the former Soviet Union evolved into a monolith. The management of the huge chain of research institutes was accomplished by a tiny but extremely influential élite of the presidium of the Russian Academy of Science (RAN). Whereas virtually all other institutions of the Communist era are now eroded, the academy remains basically unchanged. This conservatism is one of the main reasons for the transformation of a scientific 'brain drain' into an exodus, a phenomenon without any historical precedent. Members of the RAN élite are facing the prospect of becoming generals without an army. The devastating state of basic science has started to attract the attention of the government and of the Russian parliament (state *duma*). Many consider this to be the last chance for the survival of Russian science; some believe it is already too late.

Like almost everything in Russia today, science has become vulnerable to political manoeuvring. Views about the future of science in Russia tend to be pessimistic or optimistic, depending on political bias. Pessimists include the presidium of RAN, an extremely conservative agency resisting any change in its status; the academy's professional unions, mostly represented by union bosses left behind by the former Soviet system and who still openly support Communism; and the parliamentary faction of 'Russian Communists' — all these groups exemplify organizations that emphasize the disastrous situation for scientists and the government's unforgivable inability to meet their needs. Their forecasts are gloomy and are invariably summarized by the assertion that if the science budget is not increased today, tomorrow's science will die. On the other hand, the Democrats, represented by the minister of science, Boris Saltykov, and the chairman of the *duma* subcommittee on science, Nikolai Vorontsov, although not denying the hardships, are optimistic.

Both Saltykov and Vorontsov talk positively of the changes that have been generated in the Russian scientific community by the creation of the International Science Foundation (ISF) and the Russian Foundation for Basic Research (RFBR). The establishment of a system of state research centres and other organizations that are individually financed has also had a positive effect. Both men also

concede that scientists whose work is of no international value and researchers who try to secure financing on the basis of their administrative position or connections, as opposed to scientific merit, will become the first victims of the crisis. Meanwhile, grants, joint international projects and other types of money-earning activities will enable most scientists whose work is recognized outside Russia to survive.

Vorontsov suggests the following biological analogy for what has happened and is happening in Russian science. He compares its past to a strong monotype lumber forest that suppresses any new growth at its roots. When the forest is felled, new forms of life begin to sprout in its place. Many will not survive, and there is even a danger that the land will become a swamp, but the monotype is gone.

Saltykov hopes that Russian science today has finally reached the bottom of the financial chasm into which it has been falling at an ever increasing rate during recent years, and that a slow budgetary inflow will start this autumn. (ISF, following administrative delays, has finally started distributing grants.) He sees the current financial situation as critical and believes that unless the scientific budget is rethought in the early autumn, "a serious drama" will ensue. Nevertheless, Saltykov is optimistic given the first small victories won in 1994 in the battle for the increase of state allocations for science. The decrease in inflation and relative stabilization of the financial market allow him to hope for more significant increases.

There is not one official in the Russian

administration who advocates the need to reduce the scientific budget or to leave it at its current level. In addition, the Russian premier, Victor Chernomyrdin, who promised to allocate \$12.5 million to the ISF programme, despite all expectations for the worst, has fulfilled his promise.

Vorontsov also is expecting several financial breakthroughs. "Our subcommittee has already succeeded in increasing the budgetary allocations for basic research by almost 100 per cent", he says. "We are determined to spend this money on a competitive basis." He is determined to establish special foundations for channelling all additional funds obtained for science through parliament.

Even if state financial support is augmented, it will not be able to meet all of today's needs. "More money will be available", says Saltykov. "But in the near future there will not be significantly more money. We are not capable of supporting everything that was accumulated in the Soviet Union, and we do not need to." He is much more concerned about structural reorganization. The ministry's actions aimed at achieving this goal are not limited to the establishment of the RFBR and state research centres. Saltykov says that soon a foundation for the support of the humanitarian sciences will be created, and it will be organized similarly to the RFBR.

Yet there have been serious failures during the past year. Saltykov's concept of state professorships was torpedoed last autumn by RAN's president, Yuri Osipov. At the time, Osipov, who felt

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The academy — still exerting iron control over Russian scientists.

threatened by the possibility of several thousand Russian scientists working for a monthly salary of \$1,000 and the inevitable subsequent decline of RAN prestige, persuaded President Yeltsin to issue a decree establishing the title of "outstanding scientist", a step between a doctor of science and a corresponding member of the academy. These scientists were to receive a stipend of 75,000 rubles a month (about \$30 at the current exchange rate). But even this idea turned out to be stillborn: many of the appointed outstanding scientists are today receiving no stipends at all because an appropriate specification was not made in the budget; consequently, the responsibility for the stipends was transferred to the institutes with their extremely limited funds.

Meanwhile, the exodus of scientists which Saltykov was hoping to stop by creating stipends of \$1,000 per month, continues to accelerate. According to data from the Ministry of Sciences, up to 80 per cent of Russia's theoretical mathematicians have already left the country, for example. More than half of the researchers from the highly rated Landau Institute of Theoretical Physics have also gone. Many believe that the critical mass of researchers in many key fields is now lost. Still, the 'state professorship' concept has not been abandoned altogether, although it has certainly changed shape over the past year. Now the funds are intended to support young scientists under the age of 40. According to Saltykov, Chernomyrdin supports the idea.

Money is still critical. Scientists' salaries, which are mostly below the survival minimum, are often not paid for several months. This has resulted in an average monthly salary for the first quarter of 1994 of roughly 30,000 rubles — a sum that can be stretched to last for a week at most.

Of course, this does not mean that every scientist is on the brink of starvation. Like all state employees, most are compelled to look for income elsewhere — only a few lucky ones find jobs connected to their own profession. The exodus to the West (where at least the people concerned continue to contribute to the international scientific effort) and to business (those who choose this course probably should have left science anyway) is augmented by another hidden drain: a person who continues to be employed as a researcher but is incapable of making a serious contribution because his or her attention is diverted by the need to supplement his or her salary. Nobody knows the real dimensions of this internal 'brain drain'.

Corruption is another factor that makes the situation even worse. The directors of many institutes involved in basic and applied research make fortunes for themselves renting out the institutes' space and investing money intended for salaries. Their own, extremely high salaries are

fixed, whereas researchers are forced to take unpaid leave or to continue with their experiments without receiving a penny. Some directors intentionally force their employees to leave in order to free up some more space for rent. Such methods have been described by a deputy of the Moscow City *duma*, Dmitry Katayev, and Saltykov acknowledges the problem, although without specifying names. Saltykov is powerless to do anything to rectify the situation because he still has no authority over RAN's business.

The truth is that despite the uncertain future and shaky present of science in Russia, over recent years the system has not suffered from any radical external changes. All the institutes still exist and even though life in them often seems in danger of being lost, scientists continue to receive their salaries (eventually) and to make trips abroad. The directors and other officials continue to hold their positions and even regularly report to RAN's leadership about spectacular achievements associated with their institutes. These reports are often impressive because many researchers who actually work abroad retain their Russian affiliations. For example, a researcher from one of the leading RAN institutes who has been working in a US research university for a couple of years, visited his parents in Moscow this summer. During a short visit to the institute, he was pleasantly surprised to find his name on the "Board of Honour" as the most prolific scientist there. He then accidentally met the director. It seemed that the director had no idea that his best employee had not actually worked in the institute for almost 2 years. Nobody knows how many such 'dead souls' decorate the facades of RAN institutes. Long before Gogol's Chichikov, corrupted bureaucracy in Russia was skilful in hiding miserable reality behind fake facades. Russians even have a term for this sort of deception: "Potyomkin's village" (after Count Potyomkin, a favourite of Catherine the Great, who impressed his contemporaries by creating entire fake villages to deceive the tsarina). It is possible that we have witnessed an even more impressive achievement: the entire academy, with its many institutes, has become a huge 'Potyomkin's village'.

If this is true, one can easily see the reason. Although back in 1992 analysts were already asserting that the state of science in Russia was at the "point of bifurcation", nothing serious has happened since. Even the Serpukhov particle accelerator, the project the government cannot afford, is still formally alive. If the tunnel for the accelerator were finished, it could be conserved until better times. About 180 metres of the 23 kilometres remain to be completed, and if the construction were halted now the tunnel would be irreversibly flooded. Nobody in

the government will take the responsibility for the decision about how to proceed. Is scientific research in Russia going to hang up at the 'bifurcation point' (or in the form of a giant Potyomkin's village) until the science itself completely dissipates? Although this is still its most probable fate, new political manoeuvring provides some chance of the alternative, more favourable outcome.

A few weeks ago the state *duma* started preparations for the passage of the law on science and technology policy: three drafts are now being reviewed. One, submitted by the Federal Assembly, reflects the point of view of the 'Russian Communists' faction and the deputies who support the president of RAN. Another, submitted by the government, was developed by the Ministry of Science. The third, submitted by the "YABLOKO" *duma* faction, was prepared by Alexei Zakharov, the secretary of the *duma* subcommittee on science.

The first draft is the subject of an intense discussion initiated by one of its authors, Alexander Shevelukha. All three drafts were recently discussed at a special round-table conference in St Petersburg. A surprisingly large number of St Petersburg's scientists took part. Most supported the YABLOKO draft, but some believed that it should be integrated with the government's draft.

At the initiative of Vorontsov, all three drafts will be sent to the hundred largest institutes throughout Russia for discussion and review. An expert commission, chaired by the *duma* deputy, Anatoly Shabad, will review all opinions and make priority recommendations.

Zakharov, one of the main organizers of the 1991 RAN Conference of Scientists (the conference's proposal to deprive the president of RAN of absolute power over the basic research budget was turned down), believes that his draft contains an opportunity for scientists to achieve the goal that the conference failed to reach: to free themselves from RAN's corrupted bureaucracy. "This is the last chance for our scientists to participate in the making of their own fate", he says.

In recent months, the interests of scientists in "making of their own fate" have risen sharply. Whatever the phrase might mean, it would be naive to think that their voices, even if they are acknowledged by a special parliamentary commission, will have a significant influence on the distribution of power. That would be possible only if one of the main opposing forces, either the Communists or the Democrats, could use the voices of scientists to upset the established balance of power. □

Vladimir Pokrovsky is a freelance science journalist who regularly contributes to *Nature's News* section.

Punctuated equilibrium by computer

Can a mathematical model of evolution dealing with species rather than their members as the substrate of natural selection accurately represent reality? Probably not, but it is suggestive and stimulating.

THAT physicists are itching to take over biology is now well attested. People are forever suggesting explanations for the fine details of cellular behaviour, from the ratchet-like behaviour of muscle fibres to the self-assembly of microtubules and even the tendency of lipid bilayer membranes to form cylinders with μm dimensions. But surely only a brave physicist would take on Darwin on his home ground, the theory of evolution, let alone Gould and Eldridge on punctuated equilibrium, at this early stage in the take-over game. That might be everybody's expectation. Predictably, it has now been falsified.

The guiding spirit behind the latest flurry of activity appears to be Per Bak, at the Brookhaven National Laboratory. Some years ago, with C. Tang and K. Wiesenfeld, he offered the world the sandpile as a model of what he called the self-organized critical behaviour of a dynamical system (*Phys. Rev. Lett.* **58**, 381; 1987). The point of that model is that it represents a system far from equilibrium (which would be a layer of sand with a horizontal flat surface). If the sandpile is in the critical condition, the addition of a few grains of sand (in principle, just one of them) to the apex will trigger the formation of avalanches on the sides of the pile until a new position of equilibrium, perhaps metastable, is reached.

What has that to do with evolution, punctuated or otherwise? Well, may it not be reasonable to think of the long periods of evolutionary stasis as the times when sandpiles are slowly growing, and the avalanches triggered in a flurry at criticality as the bursts of speciation events that show up in the fossil record...? Whatever the truth of the matter, there has been a minor avalanche of articles on the theme since last December, all of which acknowledge their inspiration by Stuart Kauffman, the physician turned theoretical biologist at the Sante Fé Institute, whose book *The Origins of Order* was haughtily panned last year (*Nature* **365**, 704; 1993).

The sandpile, of course, is not a good model for evolution. Bak and Kim Sneppen from the Niels Bohr Institute, explained that at the end of last year (*Phys. Rev. Lett.* **71**, 4083–4086; 1993). Instead, they use points along the length of a line to represent the species in an entire ecosystem. This leads, the authors acknowledge, to the crudest of all possible models. A species is, after all, a collection of its individual members upon whom alone natural selection exerts its evolutionary influence. Representing an entire

species by a single entity and ascribing to it a "fitness" which is a scalar number, which is what Bak and Sneppen did, will drive many evolutionists to despair.

Bak and Sneppen talked their way around that difficulty last year. The idea is that a species in a reasonably comfortable (or stable) condition may be likened to a complex physical system with many metastable states. The common physical analogy is that of a spin-glass, essentially a randomly distributed system of magnetic spins interacting with at least their neighbours and allowed to find their lowest energy.

The possible states of such a system include at least an infinite number in which magnetized patches of arbitrary orientation are embedded in each other. Each of these is a candidate for stability relative to the states obtained by reversing the direction of an arbitrary spin in the spin-glass as a whole, but that is no assurance that the energy-minimum is the lowest of all possible. For what it is worth, the same analogy is used (at least in words) for modelling the conditions of neural networks.

That is the underlying model for the fitness of a species, taken to be the negative of its energy. Ideally, species would remain at a local maximum of the "fitness landscape", but the occurrence of mutations with the succession of the generations might easily drive them away. Extinction might be one consequence, but otherwise their best refuge would be a neighbouring peak in the landscape with, preferably, greater fitness.

So how to carry out the arithmetic? There are N species, each represented by a point along a line (whose chief purpose is that each point should have two neighbours). Each species (point) is randomly assigned a fitness-number, chosen (without loss of generality, as mathematicians say) to be between 0 and 1. Following the analogy of the spin-glass, Bak and Sneppen go on to argue that the evolutionary barriers (the impediments to further change) will be least for those species whose fitness is represented by the smallest number, so that the fitness number is taken as a measure of the barrier to further change.

The next step is to choose the species with the smallest fitness and to reassign it another random number. To allow for the circumstance that a change in the competence of one species will often change the prospects for others, Bak and Sneppen also reassign random numbers to the two neighbouring species in the line. To allow for what may be happening independently to

other species, they also select a certain number of other points along the line and randomly reassign them random numbers.

The most interesting conclusion from the numerical results is that the initial conditions are irrelevant to the final outcome. The system settles down to one in which the barrier height (or fitness) is high for most species. If, in the reassignment of fitness, a small number appears, the species concerned will be quickly dragged back to a conditions of high fitness through a series of readjustment of the species and its neighbours. So the analogy with the sandpile applies. The wave of readjustments in one species and its neighbours is like an avalanche down the slope of the pile, whose steady state is therefore a state of criticality.

That was Bak and Sneppen last December. The same authors and Henrick Flyvbjerg of Princeton, in the same issue of the same journal (**71**, 4087–4090; 1993), described a variant of the model in which the two nearest neighbours of the least fit species were not modified at each time step, but a fixed number were. (That is equivalent to the assumption that a change in the fitness of one species affects a fixed number of others chosen randomly — the line is unnecessary.) That, with a little approximation, can be dealt with analytically. The results confirm what the simulations show. And that argument has now been carried further by Jan de Boer from the State University of New York at Stonybrook and a group of associates (*Phys. Rev. Lett.* **73**, 906–909; 1994) to show that the evolutionary avalanches are just what Gould and Eldridge need to punctuate evolutionary equilibrium.

How convincing is all this? There is no doubt that the authors have found a model that reproduces punctuated equilibrium, but whether it represents the real world is less obvious. One drawback is that it allows neither for the extinction of species nor for their divergence into two or more. Another is that there is no telling whether the real evolutionary landscape is like that supposed in the model. (There are one or two references to "coordinated mutation" to account for the evolution of structures such as wings that would raise some biologists' hackles.) But the authors are right to say that previous treatments of self-organized criticality have shown these systems to be robust against detailed changes of the model. If a simple problem can be handled, a more complicated one may well give the same result. So the exercise should not be forgotten.

John Maddox

Peptide partners call the tune

Antonio Lanzavecchia and Colin Watts

EXPERIMENTS *in vitro* can tell us a great deal, but invariably a fuller picture emerges from work with living cells. That is one of the reminders to be had from the paper by Nelson and colleagues on page 250 of this issue¹. The question the authors have tackled is that of the stability of the marriage between antigenic peptides and class II molecules of the major histocompatibility complex (MHC). They find that different peptide–class II complexes have different lifetimes depending on the peptide concerned, a result which provides new avenues by which to understand the basis of antigenicity.

MHC molecules are glycoproteins dedicated to capturing, transporting and displaying peptides on cell surfaces for scrutiny by T lymphocytes. A breakthrough in understanding their function came with the demonstration that purified class II molecules could bind peptides known to specify T-cell determinants in an allele-specific manner². Affinity measurements by equilibrium-dialysis and gel-filtration techniques used in these and other early studies^{2,3} implied the existence of a receptor–ligand equilibrium that might be influenced by competing ligands (that is, other peptides) according to the laws of mass action. In other words, one peptide might be exchanged for another provided its affinity and concentration was high enough.

The concept of peptide exchange seemed to offer flexibility and economy, in the sense that class II molecules might be reused like other receptors. It was pursued enthusiastically by the immunological community, not least because peptide exchange on pre-formed complexes would potentially allow displacement of self-peptides involved in autoimmunity using allele-specific ligands.

Some reports — of work both *in vitro*^{4,5} and in living cells^{6,7} — were consistent with this idea. Others were not. First, it became clear that peptide acquisition by class II was an integral part of its biosynthetic itinerary and that, as demonstrated for class I MHC^{8,9}, peptides provide a vital stabilizing influence driving cell-surface expression once the invariant chain has been removed (reviewed in ref. 10). Second, earlier studies revealed a steady accumulation of antigenic determinants in antigen-presenting cells exposed to low antigen concentrations and a long half-life estimated to be at least 24 hours¹¹. Third, there were consistent reports that peptides dissociate slowly from isolated class II molecules³.

To clarify the question of peptide stability *in vivo* it was necessary to measure the stability of complexes and compare their

half-lives with those of the class II molecules to which the peptides are bound. Both functional and biochemical assays demonstrated that the lifetime of complexes was long and, more importantly, similar to that of class II molecules themselves¹². Moreover, it was found that existing complexes could not be displaced by exogenous peptides. Thus binding was essentially irreversible and MHC molecules emerged as receptors which are used in a disposable fashion.

Nelson *et al.*¹ now extend this type of study and show that the lifetime of a class II molecule depends on the 'quality' of the peptide bound. They have used different peptides contained within or related to residues 46–61 of hen egg lysozyme (HEL), which correspond to a well-characterized T-cell determinant that binds to the mouse class II molecule I-A^k. Several thought-provoking observations are made.

They first show that peptide–class II complexes have a long half-life, and that this half-life is not decreased even in the presence of a large excess of peptides which are actually better binders. This contrasts with reports suggesting short half-lives and peptide displacement in the same system^{6,7}, but is similar to results obtained in the human system¹². The second — and striking — finding is that the lifetime of the complexes depends on the length and sequence of the peptide, and correlates with the capacity of the peptide to stabilize class II molecules in the presence of SDS. Peptides which stabilize a higher proportion of class II molecules in SDS confer 2–3 times longer lifetimes than those that only partially stabilize.

Nonetheless it is now clear from this and other studies^{12,13} that SDS-unstable molecules are not necessarily empty, and that a single peptide species can be found in both stable and unstable populations. Interestingly, Nelson *et al.* indicate that peptides that show relatively short half-lives and confer low SDS stability do not necessarily show faster 'off rates' from class II molecules *in vitro*. This implies that cells can discriminate between stable and unstable complexes and impose a degree of 'quality control' analogous to that seen in other cellular compartments¹⁴. The SDS-unstable yet peptide-occupied molecules described by Nelson *et al.* seem to be different to those generated from soluble class II molecules by Sadegh-Nasseri *et al.*¹⁵; in that instance, the SDS-unstable form dissociated peptides within minutes and may represent an intermediate in the assembly of class II–peptide complexes that would be

difficult to detect in living cells.

A corollary of the findings of Nelson *et al.* is that it should be possible to prolong the half-life of class II molecules by providing better, naturally processed peptides. This is elegantly shown by demonstrating that incubation of cells with HEL can effectively extend the lifetime of the stable class II molecules by contributing peptides that stabilize better than the endogenous ones. HEL also increases the fraction of class II MHC that mature¹⁶, so both yield and lifetime are increased by provision of stabilizing peptides.

What is the basis of complex stability? The capacity to stabilize and confer a long half-life is likely to correlate with the number and quality of peptide–MHC contacts made. In the human class II DR1/HA 306–318 structure reported by Stern *et al.*¹⁷, 12-peptide residues make contact with class II MHC. Perhaps it is not surprising that in the present experiments short (ten amino acids) peptides or analogues that may have lost an anchor residue do not stabilize well in SDS and confer short half-lives.

How might quality control operate? Nelson *et al.* indicate that, for a given peptide/I-A^k complex, the ratio of SDS-stable to SDS-unstable molecules surprisingly remains constant during the chase, suggesting that SDS-stable and -unstable forms may interconvert in living cells. Because complexes with a high unstable-to-stable ratio have shorter half-lives, cells may have some means of preferentially removing molecules that 'read out' as SDS-unstable *in vitro*, either because they release peptides or because there is a subtle difference in structure.

Originally, peptide stability was perceived as a problem for antigen loading

- Nelson, C. A., Petzold, S. J. & Unanue, E. R. *Nature* **371**, 250–252 (1994).
- Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Unanue, E. R. *Nature* **317**, 359–361 (1985).
- Buus, S., Sette, A., Colon, S. M., Jenis, D. & Grey, H. M. *Cell* **47**, 1071–1077 (1986).
- Pedrazzini, T., Sette, A., Albertson, M. & Grey, H. M. *J. Immunol.* **146**, 3496–3501 (1991).
- De Kroon, A. & McConnell, H. M. *J. Immunol.* **152**, 609–619 (1994).
- Adorini, L., Appella, E., Doria, G., Cardinaux, F. & Nagy, Z. A. *Nature* **342**, 800–804 (1989).
- Harding, C. V., Roof, R. W. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4230–4234 (1989).
- Townsend, A. *et al.* *Cell* **62**, 285–295 (1990).
- Schumacher, T. N. *et al.* *Cell* **62**, 563–567 (1990).
- Cresswell, P. A. *Rev. Immunol.* **12**, 259–294 (1994).
- Lanzavecchia, A. *Immun. Rev.* **99**, 39–51 (1987).
- Lanzavecchia, A., Reid, P. A. & Watts, C. *Nature* **357**, 249–252 (1992).
- Germain, R. N. & Rinker, A. G. *Jr Nature* **363**, 725–728 (1993).
- de Silva, A., Balch, W. E. & Helenius, A. *J. Cell Biol.* **111**, 857–866 (1990).
- Sadegh-Nasseri, S., Stern, L. J., Wiley, D. C. & Germain, R. N. *Nature* **370**, 647–650 (1994).
- Germain, R. N. & Hendrix, L. R. *Nature* **353**, 134–139 (1991).
- Stern, L. J., Brown, J. H., Jardetzky, T. S. & Wiley, D. C. *Nature* **368**, 215–221 (1994).
- Neefjes, J. J., Stollorz, V., Peters, P. J., Geuze, H. J. & Ploegh, H. L. *Cell* **61**, 171–183 (1990).
- Davidson, H. W., Reid, P. A., Lanzavecchia, A. & Watts, C. *Cell* **67**, 105–116 (1991).
- Fairchild, P. J., Wildgoose, R., Atherton, E., Webb, S. & Wraith, D. C. *Int. Immunol.* **5**, 1151–1158 (1993).

because class II was imagined to be pre-saturated with self-peptides — hence the search for an exchange mechanism operating in living cells. However this would only allow transient appearance of antigenic peptides which would then be replaced by self-peptides during further rounds of exchange, drastically limiting the half-life of antigen. In fact peptide stability now emerges as a key feature in antigen presentation. Loading of processed peptides on newly synthesized class II molecules^{18,19}, and selection for the most stable complexes, allow antigen-presenting cells to accumulate a number of complexes sufficient to trigger a T cell. Such long-lived complexes are likely to

dominate in both the establishment of tolerance²⁰ and in the antigen-driven immune response.

In all, then, the common theme of receptor reuse and recycling does not seem to apply to MHC molecules. The data point to an essentially monogamous relationship of MHC with peptides that best serves the physiological requirements of the immune response. □

Antonio Lanzavecchia is at the Basel Institute for Immunology, Grenzacherstrasse 487, CH 4005 Basel, Switzerland. Colin Watts is in the Department of Biochemistry, University of Dundee, Dundee DD1 4HN, UK.

SUPERNOVA 1994I

Violent events in M51

Robert P. Kirshner

BRIGHT supernovae stimulate bright ideas. On page 227 of this issue, Ken Nomoto and his collaborators present a possible history for the star that became supernova 1994I, the brightest such event so far this year. Since early April, observers have been delighted by supernova 1994I, located in the nearby Whirlpool Galaxy (M51 or NGC5194, depending on catalogue). Reaching magnitude 13, this was an easy target for small telescopes, and a worthwhile one for the target-of-opportunity programme on the Hubble Space Telescope. Nomoto *et al.* have translated the sketchy early observations into a plausible picture for the evolution and explosion of a binary star which underwent unusual and thorough mass transfer.

The spectral classification of supernovae carries a distasteful aura of botany for many astrophysicists. We want to develop a straightforward classification that reflects the underlying mechanism of the explosion. But the observations come first and the understanding later, so that the classification reflects the observable properties of the optical spectrum lines visible near maximum light. This is a practical way to work: for many supernovae, there is only one spectrum obtained shortly after discovery. Making sense of the spectra, and sorting them into groups that are broad enough to include variations on a theme while having sharp enough boundaries to separate events that come from different stellar histories and different explosion physics, is an uncertain business.

In the beginning, there was Zwicky. Fritz Zwicky was the pioneer of supernovae studies, instigating a search for supernovae in the 1930s at the then undeveloped site of Palomar Mountain using the 18-inch Schmidt telescope. Zwicky

found that the supernovae were quite homogeneous, although their spectra were enigmatic, showing no lines from hydrogen, the most abundant element in the Universe. In 1941, Minkowski made the important discovery that some supernovae show strong hydrogen lines, and are

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Bright spark — supernova 1994I explodes in the Whirlpool galaxy.

distinctly different from the ones studied previously. So Zwicky's type of supernovae became type I and Minkowski's became type II. The conventional physical picture for these identifies type I with the sudden thermonuclear explosion of a carbon- and oxygen-rich white dwarf, which has lost its hydrogen envelope, and is pushed to ignition by mass transferred from a binary companion. The theory for type II has a core collapse in a massive star, leading to the formation of a neutron star, copious emission of neutrinos, and the disruption of a giant star, as observed for the century's great supernova, SN1987A in the Large Magellanic Cloud.

This more-or-less sensible division into types I and II has not stood the test of time. There are now three established

varieties of type I, imaginatively dubbed Ia, Ib and Ic, with well observed prototype events and significant differences in their spectra at maximum. None has hydrogen, but type Ib supernovae show strong helium lines and do not show the characteristic absorption from Si II at 6,150 Å seen in type Ia. The type Ic supernovae show neither hydrogen, nor the characteristic silicon of Ia, nor strong helium lines. The type II supernovae may also need subdivision, as suggested by last year's Supernova of the Year, SN1993J (see P. Murdoch, *Nature News & Views* **363**, 668; 1993), but that is another story. Supernova 1994I seems to belong to the Ic category.

Is this classification superficial or deep? The evidence is accumulating that supernovae of types Ib and Ic are fundamentally different from Ia: most probably they result from core collapse in a star that was once massive, but has lost its entire outer envelope. What Nomoto *et al.* have done is to construct a possible evolutionary history for type Ic supernovae that would lead to a core that was ready to collapse with an atmosphere containing no hydrogen or helium. It takes some sleight-of-hand, transferring the atmosphere to a binary companion; nonetheless, Nomoto *et al.* offer not one but three possible histories for this prestidigitation. By considering the frequency with which various initial conditions might arise, they show that type Ic supernovae should be rarer than type II, which is a testable prediction.

More concretely, they also compare the theoretical affects of exploding one of these stellar nubbins (about 2 solar masses of carbon and oxygen left from a star of 6 or more solar masses) with the observations of SN1994I. Their core-collapse model is not very detailed, but it agrees with the rapid decline of the flux as observed in the initial decay, which they attribute to the small amount of mass in the envelope. The tail of emission that extends for months after the explosion is powered by the radioactive decay of ⁵⁶Ni produced in the explosion. To match the observed flux, they need about 0.08 solar masses, which is comparable to the amount produced in other core-collapse supernovae (including type IIs).

The comparison of supernova light curves with the predictions of the model seems satisfactory as far as it goes. A more stringent test for the picture presented by Nomoto *et al.* will come from comparing the predicted spectrum with the observations. Because the spectrum includes information about the chemical composition as a function of velocity, it provides a powerful diagnostic for supernova models. You can get the light curve right without matching the spectrum — in which case, it's back to the drawing board.

Supernova classification looks simple, or even simple-minded, but when used

properly, it can help us toward a deeper understanding of the events that create supernovae. And as supernovae in turn are the events that create the elements out of which we are made, it seems worthwhile to get the facts straight. Nomoto *et al.* have taken a step in the direction of understanding SN1994I, and we have a rich observing season ahead in which the

debris of the supernova will become transparent, to provide another observational chance to test this model for the origin of type Ic supernovae. □

Robert P. Kirshner is in the Harvard-Smithsonian Center for Astrophysics, Harvard University, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

EVOLUTIONARY BIOLOGY

Parents prefer pretty plumage

Mark Pagel

WOULD parents, if given the chance, choose among their offspring on the basis of a gaudy ornament? It would appear so. On page 240 of this issue¹, Lyon *et al.* describe experiments with the American coot (*Fulica americana*) in which parents preferentially feed those chicks with colourful plumage over those lacking such plumage. This is reputedly the first experimental evidence that parents develop preferences among their offspring on the basis of an ornament, and provides a new impetus for the study of natal colouring.

Adult American coots are greyish-black birds with a spot of white on their bills. Coot chicks, however, sport bright orange plumes (see cover) that contrast sharply against their dark natal down. They also have a bald pate (hence the expression 'bald as a coot') that can turn bright red. Coot chicks do not beg vocally for food, but rather beg by displaying, in competition with other chicks in the clutch, their orange plumage to their parents.

Lyon *et al.* compared the feeding rates, and the growth and survival rates, of chicks whose orange plumes had either been trimmed or left intact. In black and orange groups, respectively, either all or none of the chicks in a clutch had their orange plumage removed. Surprisingly, chicks in these two groups did not differ on any of the three outcome measures. In experimental groups, half of the chicks were trimmed and half left untrimmed. In these mixed broods, orange chicks were fed at a higher rate, grew more rapidly, and had higher rates of survival than their trimmed clutch-mates, but were broadly similar to chicks in the two control broods. So coot parents fancy their young ornamented, but only when they have a choice between ornamented and non-ornamented offspring. Why should this be?

Manipulating only the experimental factor of interest — here the chicks' colour — can be fiendishly difficult in behavioural ecological studies. Trimming the chicks may alter their behaviour, or their viability, or even make them appear smaller, as well as altering their colour. But the absence of differences between the black

and orange groups suggests that, in itself, trimming did not affect viability, and the authors assert that trimmed chicks did not behave differently. Coot parents may also tend to feed smaller chicks more, not less. One might also quibble that parents treated orange versus black chicks differently because their experience with these two colours differs². The authors may be on safe ground here because older chicks and adult coots are black.

The experiments of Lyon *et al.* do not explore the possibility that coot parents discriminate among degrees of ornamentation in their young. But, if parents do, at least three sorts of explanation might be offered for the authors' results. If the cost of the orange plumes (energetic costs of producing them plus increased risk of predation) mean that only the fittest of chicks can produce them, then parents should preferentially feed well-ornamented chicks. Alternatively, the plumes and parents' preference for them may have co-evolved via a 'runaway' or Fisherian³ process in which an initially advantageous plumage trait becomes exaggerated because the parents prefer it. A subset of this explanation has the plumage trait evolving to a fixed and pre-existing parental preference for it. However, whether the Fisherian process works for other traits than those that are sexually selected has not been investigated.

A third explanation rests on the coots' practice of brood reduction. Eggs hatch over a series of days, and, on average, later-hatched offspring suffer relative neglect and have higher mortality. Yet orange chicks in the experiments did not suffer higher mortality when they were late in the brood, whereas black chicks did. Coots' natal colouring normally disappears by three weeks after hatching. Parents may, therefore, use the degree of colouring to assess age; black chicks in the experiments were, perhaps, regarded as old enough to look after themselves.

If coot parents do use the natal plumes to assess age, coot chicks should use their plumes to try to convince their parents that they are young. Presumably, lying about one's age would be kept in check by

the increased time-span over which a chick had to pay the costs of producing or having the ornament, and the possibly decreasing benefit of parental feeding to older chicks. Were this the case, parents should feed orange chicks more in the presence of black chicks, but they should not neglect their chicks as the whole brood turns black near the time of fledging — a pattern consistent with the authors' results.

Coot chicks also use their bald pate to beg for food. Why should coot chicks employ (at least) two begging signals? Two signals may each provide a different piece of information, or two signals together may provide a more reliable estimate of the same quality. Alternatively, one signal may indicate something about the signaller's quality, whereas the other may have evolved via the Fisherian process. Theoretical studies^{4,5} suggest that adults can have more than one trait for attracting members of the opposite sex — for example, a male bird may have a long, colourful tail and a head plume. Coots are interesting in this regard because one of their two signals may provide information about an ephemeral condition (the colour of their bald pate may signal current nutritional or disease status, for example) whereas their plumes may signal something more immanent. If so, this would help to integrate a contrasting view of begging in chicks⁶, that the degree or strength of begging is an honest indicator of offspring need, with parents preferentially feeding the most needy, and not necessarily the most fit.

A tentative explanation, then, for why American coot parents choose among their young on the basis of a natal ornament is that, for brood-reducing species, the accurate assessment of which offspring to neglect may be under strong selection. Chicks, in turn, produce the ornament, and possibly other signals, to avoid being among the dispossessed. Whatever the eventual interpretation of the work by Lyon *et al.*, the idea of parental preference for young on the basis of ornaments should stand astride the conventional views of natal colouring as attempts at crypsis, or as sign-stimuli or species-recognition signals. Evidence that parents prefer young with symmetrical ornaments⁷ may be just around the corner. □

Mark Pagel is in the BBSRC-NERC Ecology and Behaviour Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

1. Lyon, B. E., Eadie, J. M. & Hamilton, L. D. *Nature* **371**, 240–243 (1994).

2. Dawkins, M. S. *Anim. Behav.* **19**, 575–582 (1971).

3. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).

4. Pomiankowski, A. P. & Iwasa, Y. *Proc. R. Soc. Lond.* **B253**, 173–181 (1993).

5. Iwasa, Y. & Pomiankowski, A. *Evolution* (in the press).

6. Godfray, H. C. J. *Nature* **352**, 328–330 (1991).

7. Møller, A. P. *Nature* **357**, 238–240 (1992).

Nature and nurture of myopia

Josh Wallman

THE mystery of how organs reach and maintain their characteristic shapes has personal significance for perhaps a third of the world's people — those who require spectacles because their eyes are too long for the cornea and retina to focus distant images on the retina, and are therefore myopic. Experiments, especially on chicks, have shown that what the eye sees influences how it grows. This line of research, and related work on humans, came in for detailed scrutiny at two meetings earlier this year*.

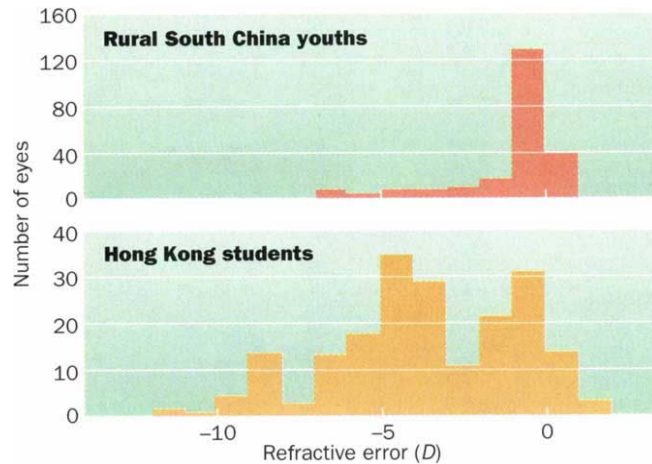
Although myopia and hyperopia (farsightedness) are generally stable conditions in adults, in newborns these refractive errors rapidly diminish resulting in emmetropia; this is the condition in which the length of the eye is well-matched to the focal length of its optics. We know that the young eye can use visual information to determine whether to grow in the direction of myopia (that is, longer) or hyperopia (shorter) because if we impose myopia or hyperopia on a chick or tree-shrew eye using spectacle lenses the optical mismatch rapidly disappears¹⁻³. (Of course, after compensating for imposed hyperopia, eyes would be myopic without the spectacles.) Now it appears that primate eyes compensate in a similar manner (L. F. Hung, Univ. Houston).

This result is of clinical importance because it hints that correcting a child's myopia with spectacles might lead to the eye compensating for the lenses and the clinician compensating for the eye's compensation — an iatrogenic positive feedback loop exacerbating the myopia. On the other hand, because lack of clear images leads to 'deprivation myopia', not correcting myopia might also be ill-advised. We need to know what visual cues the primate and chick eyes use to guide their growth, and how the two differ.

One difference seems to be that whereas chick eyes stop growing in the myopic direction within days of removal of the diffusers or negative lenses, primate eyes continue to grow in that direction for months⁴ or years (T. Tokoro, Tokyo Medical and Dental Univ.). Monkeys that were given a year of close television viewing became myopic only

subsequently (F. Shih, National Taiwan Univ. Hospital).

What visual cues might the eye use to compensate for spectacle lenses? The classic, but unsubstantiated, hypothesis is that the amount of accommodation (focusing) tells the eye whether it is hyperopic (more accommodation) or myopic (less accommodation). But chicks can still



In contrast to youths in rural areas of China, most Hong Kong students are myopic (negative refractive errors). All data are from 18–28-year-olds (data from C. Lam, Hong Kong Polytechnic).

compensate for lenses after accommodation is eliminated by brain lesions⁵ or drugs (H. Schwahn, Univ. Eye Hospital, Tübingen).

One reason for linking accommodation to myopia is that the muscarinic antagonist atropine, which blocks accommodation, has been said to halt the progression of myopia in children and in some monkeys^{6,7}. R. H. Kennedy (Univ. Texas) reported a reduction in myopia of 1 dioptre in 214 children even when measured at 20 years of age, arguing that atropine reduces myopic progression. Similarly, in children with one eye more myopic than the other, the difference between the eyes was reduced by treating the more myopic eye (S. Chew, Rockefeller Univ. and Singapore National Eye Center).

However, atropine may not act by blocking accommodation because in chicks, which lack muscarinic receptors in their ciliary muscles, it reduces both deprivation myopia^{8,9} and compensation for spectacle lenses (C. Wildsoet, Queensland Univ. Technology). So much the better, an optimist might say: atropine acts on the retina, enabling us to treat eyes with antagonists to retinal muscarinic receptors, such as pirenzepine, avoiding the side-effects of paralysed accommodation and dilated pupils. But both drugs require doses far above those needed to block

muscarinic receptors, implying that non-specific drug effects or even retinal toxicity may be involved (M. Rickers, Univ. Eye Hospital, Tübingen). Furthermore, at effective doses, both drugs reduce synthesis of extracellular matrix in the sclera, the outer layer of the eye, in chicks (D. Marzani, City Univ., New York), and atropine disrupts rabbit sclera (S. Chew), meaning that the drugs may be interfering with normal, as well as myopic, eye growth.

Because deprivation myopia can develop in eyes with severed optic nerves, local retinal mechanisms must be important. It has long been known that the level of vasoactive intestinal polypeptide (VIP) is higher in the retinas of myopic monkeys¹⁰. Now, amacrine cells in those retinas have been shown to express more VIP messenger RNA (T. Young, Harvard Medical School). Dopamine, another retinal neurotransmitter, is decreased in myopic chick eyes, and its agonist, apomorphine, partially prevents deprivation myopia in chicks and monkeys^{11,12}. Meetings of the Association for Research in Vision and Ophthalmology (ARVO) have yielded an embarrassment of additional inhibitors in chicks: this year, naloxone (enkephalinergic antagonist), reserpine (inhibitor of release of dopamine and serotonin) and pirenzepine (M1 muscarinic antagonist); last year, a VIP antagonist, and basic fibroblast growth factor (R. Seltner and B. Rohrer, Univ. Calgary; F. Schaeffel and M. Rickers, Univ. Eye Hospital, Tübingen). "Tell me", whispered someone during one session, "is there *anything* that does not block form-deprivation myopia?"

What makes the eye elongate? In deprivation myopia in a mammal, synthesis of DNA and proteoglycans is decreased (A. Reeder, Univ. Wales), as it is in the fibrous layer of chick sclera (D. Marzani and J. Wallman, unpublished data), but opposite to the changes in the cartilaginous layer of the chick sclera¹³. Furthermore, the sclera of myopic chick eyes increases latent metalloproteinase activity (J. Rada, Univ. Pittsburgh). These observations suggest that the sclera produces myopia by remodelling during growth, rather than by stretching as once thought: although tree-shrew sclera from myopic eyes does stretch more easily than normal (T. Siegwart, Univ. Alabama).

Although the research on animals shows that the growing eye can adapt remarkably to its visual circumstances, a study published earlier this year argued that human myopia is primarily genetic in origin because children with two myopic

*Fifth International Conference on Myopia, Toronto, 22–24 June 1994. Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida, 1–6 May 1994.

parents are slightly more likely to be myopic and have longer eyes than children with no myopic parents¹⁴. One upshot was that *Time* magazine announced that the gene for eye-shape had been found, and that a pharmacological cure for myopia was just around the corner¹⁵.

This impulse to prefer simple genetic explanations to complex epigenetic ones, reminiscent of the nature–nurture debate on IQ 20 years ago, was dealt something of a blow at the Toronto meeting by evidence supporting the strong association of myopia with increased education and the rapid rise in myopia upon the introduction of schooling and modern life. Among Taiwanese students and Hong Kong medical students 70–80% are myopes, compared with 20–30% in rural areas (C. Lam, Hong Kong Polytechnic; L. Lin, National Taiwan Univ. Hospital) (see figure). Furthermore, in an isolated Inuit population in Greenland, the incidence of myopia rose tenfold in 32 years (E. Goldschmidt, Copenhagen Univ. Hospital). Because myopic children stay myopic, one can infer a secular increase in the incidence of myopia from the facts that young Finnish Inuits are on average myopic, whereas middle-aged Inuits are not (H. R. Forsius, Folkhalsan Genetic Inst.), a trend that also occurs in Hong Kong (C. Lam).

In all, it is becoming clear that some kind of visual cues actively guide the growing eyes of birds and mammals towards emmetropia. These cues might account for the association of education with myopia if they cause students' eyes to grow into focus at the distance of the page, and those of a person who largely lives outdoors to grow into focus at infinity. □

Josh Wallman is in the Department of Biology, City College, City University of New York, New York, New York 10031, USA.

- Schaeffel, F., Glasser, A. & Howland, H. C. *Vision Res.* **28**, 639–657 (1988).
- Irving, E. L., Sivak, J. G. & Callender, M. G. *Ophthalm. Physiol. Opt.* **12**, 448–456 (1992).
- Siegrwart, J. T. & Norton, T. T. *Invest. Ophthalmol. Vis. Sci.* (ARVO Suppl.) **34**, 1208 (1993).
- Troilo, D. & Judge, S. J. *Vision Res.* **33**, 1311–1324 (1993).
- Schaeffel, F., Troilo, D., Wallman, J. & Howland, H. C. *Visual Neurosci.* **4**, 177–183 (1990).
- Raviola, E. & Wiesel, T. N. *New Engl. J. Med.* **312**, 1609–1615 (1985).
- Jensen, H. & Goldschmidt, E. in *Refractive Anomalies: Research and Clinical Applications* (eds Grosvenor, T. & Flom, M. C.) 371–383 (Butterworth–Heinemann, Boston, 1991).
- Stone, R. A., Lin, T. & Laties, A. M. *Exp. Eye Res.* **52**, 755–758 (1991).
- McBrien, N., Moghaddam, H. O. & Reeder, A. P. *Invest. Ophthalmol. Vis. Sci.* **34**, 205–215 (1993).
- Stone, R., Laties, A. M., Raviola, E. & Wiesel, T. N. *Proc. natn. Acad. Sci. U.S.A.* **85**, 257–260 (1988).
- Stone, R. A., Lin, T., Laties, A. M. & Iuvone, P. M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 704–706 (1989).
- Iuvone, P. M., Tigges, M., Stone, R. A., Lambert, S. & Laties, A. M. *Invest. Ophthalmol. Vis. Sci.* **32**, 1674–1677 (1991).
- Rada, J., McFarland, A. L., Cornuet, P. & Hassell, J. *Curr. Eye Res.* **11**, 767–782 (1992).
- Zadnick, K., Satariano, W. A., Mutti, D. O., Sholtz, R. I. & Adams, A. J. *J. Am. med. Ass.* **271**, 1323–1327 (1994).
- Time* p. 38, 16 May 1994.

The shapes of things to come

Philip Ball

FOR Emil Fischer, being a pioneer of natural-product synthesis was not without its problems. So heavily did the noxious odour of skatole adhere to him and his students that they are said to have had difficulty in booking hotel rooms. Fischer's legacy has overridden such tribulations, however. His most enduring contribution to molecular science is the lock-and-key principle (*Schloss–Schlüssel-Prinzip*) advanced exactly one hundred years ago to explain the mechan-

ism of enzyme action. Only this 'induced-fit' hypothesis can easily explain how interactions remote from the binding site can influence substrate association, by restricting the flexibility of the binding pocket for example. And structural rearrangements in protein–substrate complexes, relative to the free enzyme, are now clearly seen in many X-ray structures.

This sensitivity of host structure to guest binding has been a theme of supramolecular chemistry since its inception in the

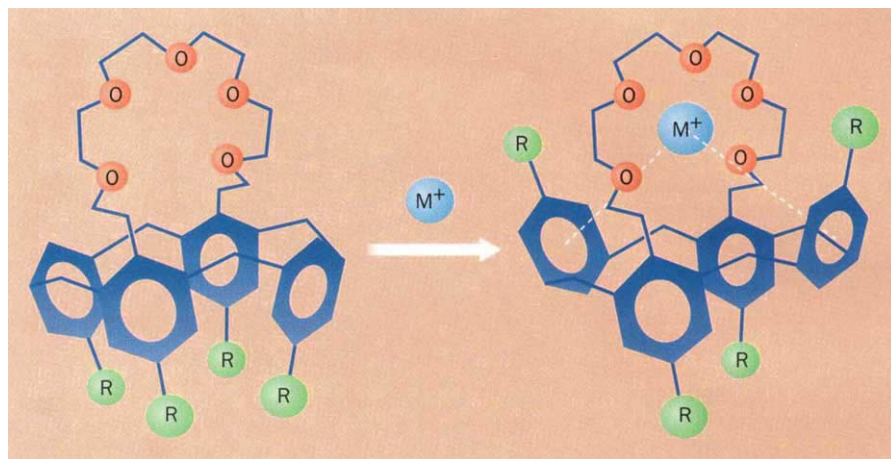


FIG. 1 Binding of metal ions by this calixarene-crown involves inversion of two of the benzene rings to allow for interactions between the ion and the π orbitals of the rings. If the substituents R are large enough (for example, isopropyl groups), inversion may be sterically favoured before the metal is bound, and this preorganization increases the binding constant.

ism of enzyme action. At a conference* marking the centenary, the strength of the analogy was most apparent from the fact that it is both precise enough and flexible enough to have survived a century of studies in chemical and biochemical molecular recognition.

Fischer's idea was that an enzyme bears a binding site that uniquely fits the shape of its substrate. Shape and geometry, then, are pictured as lying at the heart of specific molecular interactions. Today, with new protein structures appearing weekly, the importance of shape in molecular biology has never been more evident; supramolecular chemistry, meanwhile, takes as perhaps its central theme the construction of molecules with cavities that are exquisitely sculpted to allow binding of guests via noncovalent interactions.

And yet that is by no means the whole story. For example, many enzymes seem to be rather soft locks, adapting the shape of their binding sites in small but crucial ways to make a snug fit for the substrate (D. Koshland, Univ. California, Ber-

keley). Only this 'induced-fit' hypothesis can easily explain how interactions remote from the binding site can influence substrate association, by restricting the flexibility of the binding pocket for example. And structural rearrangements in protein–substrate complexes, relative to the free enzyme, are now clearly seen in many X-ray structures. This sensitivity of host structure to guest binding has been a theme of supramolecular chemistry since its inception in the 1960s, when Pedersen reported the complexation of metal ions by cyclic crown ethers. These macrocycles are conformationally labile in solution, taking on the puckered crown shape only when the ion sits in the central cavity. Binding interactions that are enhanced by host rearrangements are central to such systems as calixarene-crowns, which constitute highly selective metal binders (R. Ungaro, Univ. Parma; S. Shinkai, Univ. Kyushu) (Fig. 1) and the molecular 'clips' of R. Nolte (Univ. Nijmegen).

The balance between flexibility and rigidity is a subtle one, however. There is an entropic cost to restricting the host's (and guest's) degrees of freedom on binding (D. Williams, Univ. Cambridge), which is why preorganized, relatively rigid cavities such as those of Cram's washer-like spherands have higher binding constants for metal ions than do the floppy crown ethers.

The second new perspective on the lock-and-key principle comes from the fact that enzymes do not simply bind — they perform catalysis. Good substrate binding alone does not guarantee this function; indeed, if the effect were simply to stabilize the substrate, its reactivity

*European Research Conference on Supramolecular Chemistry, Mainz, Germany, 11–16 August 1994.

would be lowered. The true influence of the protein is rather to stabilize the transition state (TS) of the reaction relative to reactants or products.

So if supramolecular chemists are to mimic enzyme action — and this is a fundamental aim of the field — they must remember that it is the TS that counts. Often this resembles the reactants closely enough that one can design the host for binding of these latter species; but the process of catalysis is certainly enhanced if the product differs from the TS in some important way that ensures that the former is spat out instead of remaining bound as a duplex. This problem of product inhibition bedevils attempts to create oligonucleotides that can replicate via a templating mechanism (G. von Kiedrowski, Albert-Ludwigs-Univ.).

The idea of engineering TS stabilization by using TS analogues, stable molecules with structures similar to the presumed transition state, has given rise to antibody catalysis, whereby molecular catalysts are created by stimulating an immune response *in vivo* (specifically, in spleen cells) to TS analogues. The approach has now advanced to the stage where it has preparative value for traditional areas of organic chemistry such as natural-product synthesis (E. Keinan, Scripps Institute, La Jolla). Fischer would, no doubt, have been delighted.

But approaches of this sort to catalytic molecules mean that we must sacrifice the intellectual satisfaction, not to mention the practical benefit, of knowing exactly how catalysis is occurring (J. Sanders, Univ. Cambridge). The question is then: can chemists hope to synthesize catalytic hosts with the efficacy of biological counterparts? The answer is 'not yet', but the task is far from hopeless. Sanders showed that porphyrin-based molecules can accelerate a reaction to a degree only a few orders of magnitude less than typical

enzymatic reactions, an observation that can be readily rationalized in terms of TS stabilization (see L. F. Lindoy, *Nature News and Views* **368**, 96; 1994).

Supramolecular chemistry is about far more than copying nature, however attractive an aim that may be. This is evident from the interest that it is attracting from electronic engineers, physicists and materials scientists, as G. Whitesides (Harvard Univ.), who talks regularly with such people, pointed out. For supramolecular chemistry is at heart a fabrication technology, albeit at a scale far smaller than electron-beam lithography can manage.

Nowhere was this clearer than in the description of supramolecular assemblies with more components than can be counted on both hands (J.-M. Lehn, Univ. Louis Pasteur, Strasbourg). It is hard here to distinguish lock from key, yet these assemblies come together spontaneously in solution — and do so, moreover, even in the presence of molecular components that might be expected to interfere with the process. This specificity in complex molecular systems is what one finds in the cell, and Lehn predicts that such 'instructed chemistry' will take the discipline beyond the mechanical principle of lock and key to the realm of an informational science.

Information must be stored on an addressable structure, be it a linear strip such as a magnetic tape or DNA, or a lattice or grid. In the solid state, grids are most space-efficient, which is why Lehn's self-assembling molecular grid is a teasing hint of what may be possible. The grid is

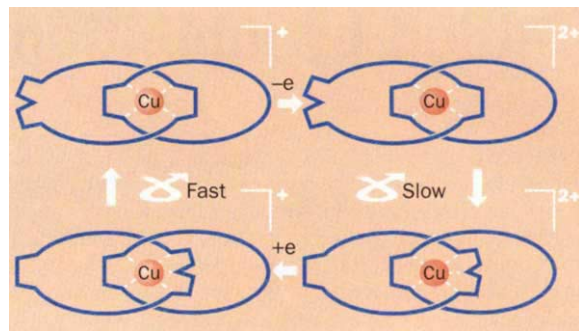


FIG. 3 Spinning the rings: electrochemical oxidation of the central Cu(I) ion in this catenane brings about a change in coordination whereby the left-hand ring flips to allow the oxidized Cu(II) species to become five-coordinate. The switching is evident from a colour change: the initial four-coordinate Cu(I) complex (top right) is deep green, but as the ring swings to produce the five-coordinate form (bottom right) — a process that takes about three days to go to completion — the colour changes to pale yellow-green. Reduction of the metal ion reinstates the original geometry.

assembled from linear rods which criss-cross at junctions bound by four-coordinate metal ions such as Ag(I). Six rods with three metal-binding sites apiece come together with nine metal ions to form a 3×3 structure (P. Baxter, Univ. Louis Pasteur, Strasbourg; Fig. 2), which is stable and can be crystallized. One might imagine that the three distinct ion sites will have different redox potentials, perhaps allowing for some degree of electrochemical addressability. The use of conducting molecular rods, or of magnetic or luminescent metal centres, suggests plentiful opportunities for these systems.

Function need not be static, and commonly is anything but. This is why mechanical molecular switches are very much in vogue. One such consists of a [2] catenane — two interlinked molecular rings — in which one ring can be flipped electrochemically between two distinct positions (J.-P. Sauvage, Univ. Louis Pasteur, Strasbourg). A copper ion is bound to the rings in the cavity where they cross, and a change in preferred coordination number of this ion, induced by electrochemical oxidation from Cu(I) to Cu(II), brings about a change in binding from one end of one ring, which is bidentate, to the other, which is tridentate (Fig. 3). As the change takes about three days to go to completion, this is hardly a practical molecular device; but the principles can surely be more widely applied.

For the future, the diversity and complexity of chemical locks will be limited only by the imagination of the locksmiths. This, suggested F. Stoddart (Univ. Birmingham) marks a progression in chemical grammar from words (molecules) to sentences (supermolecules). But what must now be learnt is the construction of paragraphs and, ultimately, the telling of stories. □

Philip Ball is an associate editor at Nature.

P. Baxter, Strasbourg

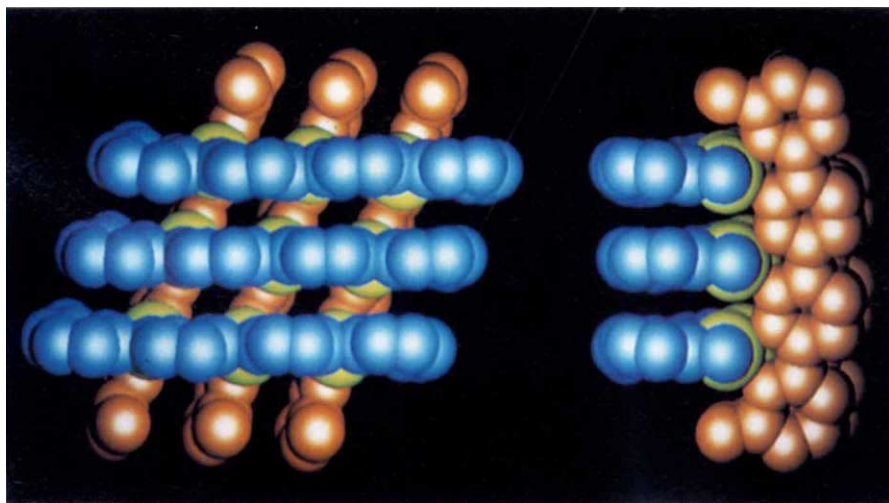


FIG. 2 Molecular noughts and crosses: this supramolecular grid, formed by complexation of six linear rods containing bipyridine units with nine silver(I) ions, self-assembles in solution. The image is derived from a crystal structure.

Stifled by inhibitions

Gordon Peters

THE division cycle of a eukaryotic cell is generally 'understood' as a series of transitions or checkpoints in which temporal order is imposed by a family of serine/threonine kinases, the cyclin-dependent kinases (CDKs), acting in concert with

ing aside the complex phosphorylation and dephosphorylation events needed to activate the various kinases, the current model invokes successive waves of CDK activity, regulated in turn by cyclins D, E, A and B (Fig. 1).

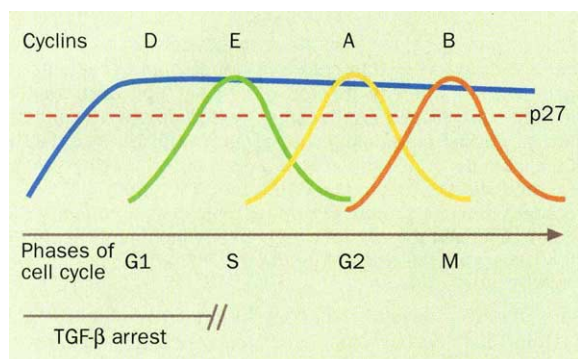


FIG. 1 The G1, S, G2 and M phases of the eukaryotic cell cycle, depicted as a linear series of events regulated by the sequential action of kinases associated with cyclins D, E, A and B. Recent findings suggest that these activities are counterbalanced by the continuous presence of kinase inhibitors, such as p27, such that 'spikes' of activity are generated when the kinase levels exceed the inhibitory threshold. In cells treated with TGF- β , E-cyclin/CDK2 activity fails to exceed the inhibitory effects of p27 and cells arrest at G1/S.

their regulatory subunits, the cyclins. Although more and more of these molecules are being found in mammalian cells, there is a comforting logic to the intimate — but not necessarily monogamous — partnerships formed between cyclins and their favourite kinases. Leav-

ing already warranted banner headlines, not least because of their connection with the tumour-suppressor genes *p53* and *Rb*. p21 is subject to transcriptional regulation by wild-type p53 (ref. 6) whereas p16 appears to accumulate in cells in which the function of Rb has been compromised, for

example in cells transformed by DNA tumour viruses^{4,10}. p16 gained further notoriety as the product of the so-called *MTS1* locus on human chromosome 9p21 (refs 11, 12), a region that is often deleted in tumour cell lines, although the significance of these deletions in primary tumours remains a matter of debate¹³.

In describing the positional cloning of *MTS1*, Kamb and colleagues¹¹ noted that the cosmid spanning the genomic locus included a DNA sequence that was highly homologous to the second exon of *MTS1*. Although the significance of this feature was at that time unclear, they named it *MTS2*, expecting it to turn out to be an independent gene. Using the p16 complementary DNA as a hybridization probe, Hannon and Beach¹ have now cloned the full coding domain of *MTS2* and find that it encodes a protein of relative molecular mass 15K, a close relative of p16, which they call INK4B. The degree of similarity ranges from about 44 per cent in the amino-terminal third to 97 per cent in the remainder of the proteins. Both molecules comprise four ankyrin-like repeats and bind directly to CDK4 and CDK6, the main catalytic partners for cyclins D1, D2 and D3. Thus, p15 and p16 appear to inhibit the function of these kinases by competing with the cyclin for binding to the kinase.

So why are there two proteins with such similar properties? The answer presumably lies in the mechanisms that regulate their expression, the remarkable feature of p15 (unlike p16) being that the levels of its messenger RNA increase some 30 times when cells are treated with TGF- β ¹. The effect is mirrored by a corresponding decrease in the CDK6-associated kinase

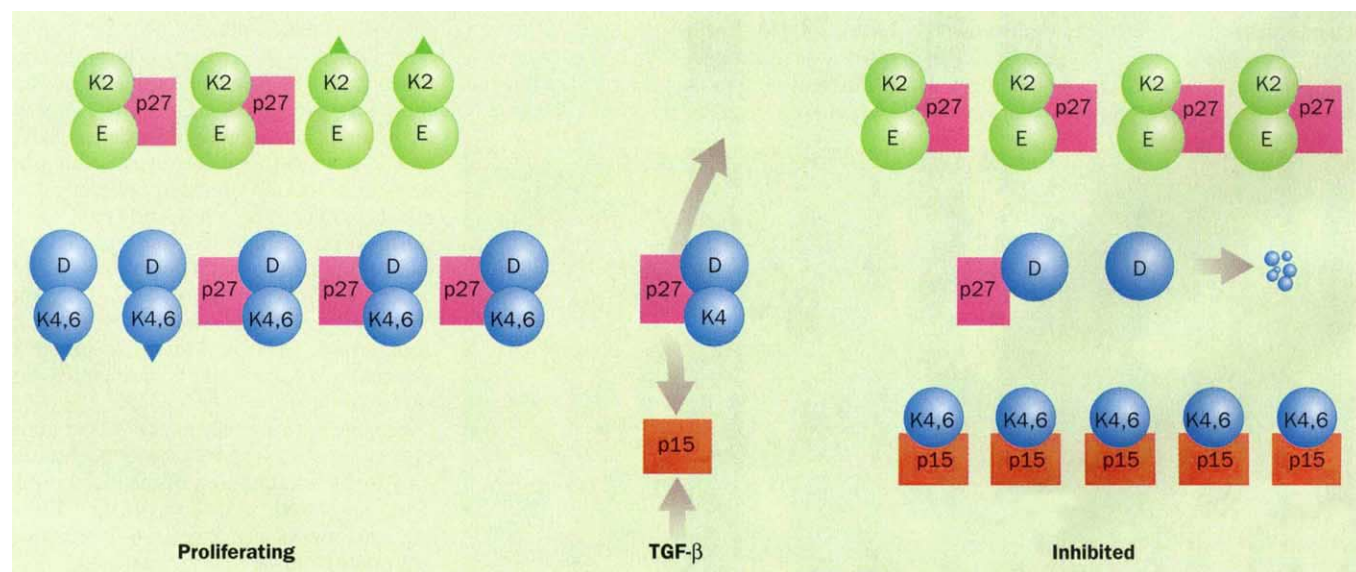


FIG. 2 In normally proliferating cells, the p27 inhibitor is associated with multiple cyclin/kinase complexes (only complexes with cyclins D and E are shown), but its partitioning leaves a significant proportion of these complexes in an active state (shown as an arrowhead on the kinase subunit). Upon treatment with transforming growth factor- β 1 (TGF- β), the distribution of p27 is altered such that E-cyclin/CDK2 (K2)

complexes are completely inhibited and cells arrest at G1/S. TGF- β also induces the expression of p15, which binds directly to CDK4 (K4) and CDK6 (K6), and destabilizes the D-cyclins. The model described in this News and Views article links these events by suggesting that the sequestration of CDK4 and CDK6 by p15 is responsible for the redistribution of p27.

activity, measured using Rb as a substrate. Because it is generally accepted that the D-cyclin/CDK complexes are at least partly responsible for the phosphorylation and functional inactivation of Rb, the induction of p15 would provide a satisfying explanation for the accumulation of hypophosphorylated Rb in TGF- β -arrested cells. But this can be only part of the story because other cyclin/CDK complexes also phosphorylate Rb.

Enter the second of the new CDK inhibitors — p27 (KIP1). It was originally identified and purified as a heat-stable protein that inhibited E-cyclin/CDK2 function in mink lung cells arrested by TGF- β or by contact inhibition¹⁴. A similar protein (p28/ICK) was observed in lovastatin-arrested HeLa cells but the identity of these proteins awaits confirmation¹⁵. In a classical reverse-genetics approach, Polyak *et al.*² used peptide-sequence information from purified p27 to isolate the corresponding cDNA. Meanwhile, Toyoshima and Hunter³ came upon the p27 cDNA in a yeast two-hybrid screen using cyclin D1 and CDK4 as bait, a similar approach to the one that yielded p16 through its association with CDK4 (ref. 4). The deduced amino-acid sequence of p27 is completely unrelated to either p15 or p16 but instead resembles p21, particularly in the amino-terminal domain. This region of the protein is essential for binding to cyclin/CDK complexes, and although the data need to be clarified it seems that p27 and p21 bind to either the cyclin or the cyclin/kinase complex and not to the CDK itself (in contrast to p15 and p16). Similarly, p27 and p21 seem more catholic in their binding properties and p27 can inhibit complexes containing cyclins A, B, D and E (refs 2, 3).

The confusing aspect of the new data is that the expression of p27 is unaffected by TGF- β , despite all the indications that it is responsible for inhibiting E-cyclin/CDK2 activity in TGF- β -arrested and contact-inhibited cells^{1,2}. Its expression is also constant throughout the cell cycle^{2,3}. So how can these facts be rationalized?

One idea is that, for cells to proceed through the G1/S transition, the level of E-cyclin/CDK2 activity must exceed the threshold of inhibition set by p27 (Fig. 1). As p27 can participate in multiple complexes, treatment with TGF- β might alter the balance of these associations and increase the availability of p27 to inhibit E-cyclin/CDK2. This would tie in with the observation that the amount of p27 associated with E-cyclin/CDK2 in proliferating cells can be increased by simply heating the cell extracts¹⁴. Although not explicitly stated in any of the papers, it would make sense if the heat-labile binding protein proposed by Polyak *et al.* were none other than D-cyclin/CDK4 or D-cyclin/CDK6 (Fig. 2). The attraction of such an idea is

that it provides a mechanism for increasing the p27 threshold in response to TGF- β . Inducing p15 expression would sequester CDK4 and CDK6, destabilize the D cyclins and release the bound p27. Of course, all this assumes that another new group of molecules will not surface to complicate things further.

No doubt more inhibitors will be identified, and there is an urgent need for some uniformity in nomenclature. A perplexing variety of acronyms has already been tried, not to mention echoes of the Strategic Defense Initiative and authors' names. In the case of p15, INK4B is hardly appropriate for a protein that inhibits both CDK4 and CDK6, and MTS2 is equally insecure until its contribution as a tumour suppressor is more fully explored. A more rational system would be arbitrary numbering in the order of discovery and the human genome mappers have already responded with *CDKN1*, *CDKN2* and so on. Such names are unlikely to prove popular in conversation, and ignore the fact that p21 and p27 are structurally

unrelated to p15 and p16, and may have quite different roles in cell-cycle regulation. Let's face it, a lot of people will be talking about these interesting proteins in the months ahead. □

Gordon Peters is at the Imperial Cancer Research Fund Laboratories, 44 Lincoln's Inn Fields, London WC2A 3PX, UK.

1. Hannon, G. J. & Beach, D. *Nature* **371**, 257–261 (1994).
2. Polyak, K. *et al.* *Cell* **78**, 59–66 (1994).
3. Toyoshima, H. & Hunter, T. *Cell* **78**, 67–74 (1994).
4. Serrano, M., Hannon, G. J. & Beach, D. *Nature* **366**, 704–707 (1993).
5. Xiong, Y. *et al.* *Nature* **366**, 701–704 (1993).
6. El-Deiry, W. S. *et al.* *Cell* **75**, 817–825 (1993).
7. Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *Cell* **75**, 805–816 (1993).
8. Gu, Y., Turk, C. W. & Morgan, D. O. *Nature* **366**, 707–710 (1993).
9. Noda, A., Ning, Y., Venable, S. F., Pereira-Smith, O. M. & Smith, J. R. *Exp. Cell Res.* **211**, 90–98 (1994).
10. Xiong, Y., Zhang, H. & Beach, D. *Genes Dev.* **7**, 1572–1583 (1993).
11. Kamb, A. *et al.* *Science* **264**, 436–440 (1994).
12. Nobori, T. *et al.* *Nature* **368**, 753–756 (1994).
13. Bonetta, L. *Nature* **370**, 180 (1994).
14. Polyak, K. *et al.* *Genes Dev.* **8**, 9–22 (1994).
15. Hengst, L., Dulic, V., Slingerland, J. M., Lees, E. & Reed, S. I. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5292–5295 (1994).

GEOCHEMISTRY

How wet is the Earth's crust?

Bruce W. D. Yardley and John W. Valley

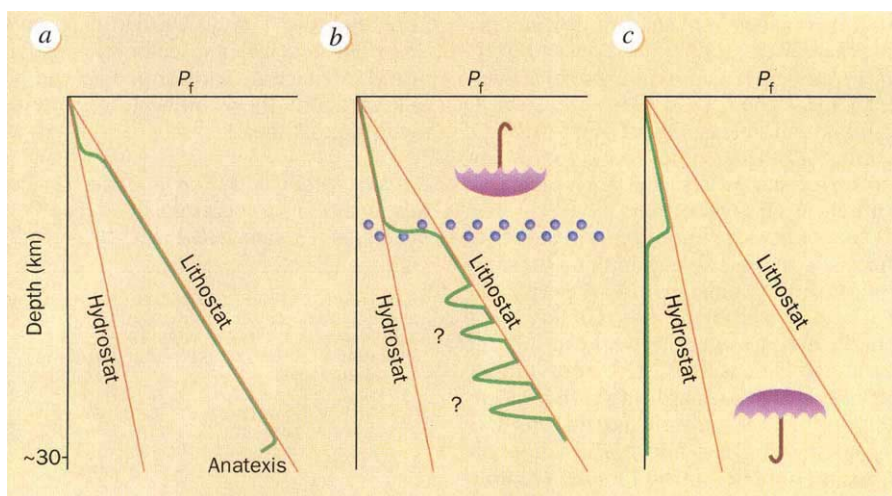
THE past twenty years have seen enormous progress in our understanding of the interior of our planet, but some very simple questions about the crust remain controversial. A key question, which has implications for our understanding of how plates deform, how ore deposits develop, and for deep waste disposal, is the extent to which fluids are present in the crust. On page 238 of this issue, R. C. Bailey¹ presents a physical model of fluid behaviour in the mid-crust which offers a resolution of some of the apparent anomalies in current ideas about crustal fluids, and also provides a sufficiently rigorous and specific treatment to discriminate between some alternative schemes.

Evidence about crustal fluids comes in three ways: studies of the present-day crust by geophysical measurements of properties such as seismic velocity, and, increasingly, by deep drilling; studies of surface rocks which originally formed in the mid- to lower crust; and laboratory studies of the conditions under which specific mineral assemblages are stable. The behaviour of fluids in deep sedimentary basins is relatively well known; the main uncertainties arise where the crystalline crust of the stable continents is concerned. A particular problem is that studies of metamorphic rocks primarily provide information about fluids during a brief and atypical period of past crustal history — active metamorphism and orogenesis — whereas geophysics and

drilling provide information about the state of crystalline crust today, when it is cooler and often shallower than during the original formation of the constituent metamorphic and igneous rocks.

Neglecting points of detail, there are two fundamentally different ways of looking at the data. One view is that the deeper, ductile crust contains an interconnected pore fluid at roughly lithostatic pressure. Alternatively, the deeper crust may be viewed as essentially dry, any fluid occurring only in isolated pores or inclusions. In the upper part of the crystalline crust, fluid occurs in fractures but is at relatively low (roughly hydrostatic) pressure. Whereas the existence of lithostatically pressured fluids is not in question for areas of active burial and metamorphism^{2–4}, the state of fluids in the stable continental crust is controversial, and the alternatives are summarized in the figure.

The evidence for interconnected deep fluids is cited by Bailey¹ and is essentially geophysical. Both seismic reflectors and zones of high electrical conductivity have been ascribed to fluid-rich zones at depth. But the interpretation is not always unambiguous. Seismic reflectors in the Siljan structure in Sweden were attributed to gas-rich zones by proponents of a mantle methane source, but have proved to be diabase sills in granite⁵, and their high conductivity is due to near-hydrostatically pressured water in fractures, not over-



Alternative models for crustal fluids illustrated by schematic plots of fluid pressure (P_f) as a function of depth. *a*, Condition during burial and progressive metamorphism of sediment to the onset of melting (anatexis). *b*, *c*, Alternatives for stable continental crust. In *b* there is a continuous fluid phase through all or part of the ductile lower crust, prevented from moving up by a zone of immiscible fluids (see page 238 of this issue)¹. In *c* the lower crust has fluid in isolated inclusions only, and is a potential sink for water from above. The umbrellas emphasize the alternative flows that the models imply.

pressured fluid. Nevertheless, the simplest explanation of increased conductivity in the deep crust is the presence of a continuous, lithostatically pressured, water-rich fluid.

A physical difficulty with this model is that the crust is probably too permeable to keep overpressured fluids bottled up at depth over geological timespans once metamorphism has ceased to replenish them. Bailey's contribution has been to show how an impermeable barrier might develop in the mid-crust, above the brittle-ductile transition, through the interference of migrating immiscible water and carbon dioxide fluids. He assumes that the fluid rising out of the deeper crust is dominated by (relatively dense) carbon dioxide which initially contains dissolved water, and as the water exsolves it plugs the conduits to CO_2 migration. This is a physically feasible mechanism to contain some types of deep fluid at high pressure. Perhaps geologists and petrologists have failed to recognize former deep 'water sills' in rocks exhumed to the surface?

The evidence for a dry deep crust is from rocks such as granulites⁶, which formed at high temperatures and are important constituents of the lower crust today. The anhydrous minerals in these rocks react very rapidly with water under the conditions that prevail in stable lower crust, and such reactions would consume all available water, possibly leaving isolated traces of salts and CO_2 (refs 6,7). Because fresh high-temperature minerals have survived in abundance to the surface, petrologists argue against the persistence of a continuous aqueous (conducting) fluid phase at depth. Where granulites have undergone retrograde hydration, this generally occurred at low temperatures along joints and fractures, after the rocks

had been uplifted through the brittle-ductile transition.

Thus the alternative to the model espoused by Bailey is that deep stable crust does not contain interconnected fluid in fluid-rich layers; any fluid is in isolated pores or inclusions, perhaps transiently migrating down along faults and shear zones^{8,9}. In this case, deep seismic reflectors are due to lithological contacts or mylonite zones, and the electrical conductivity may be due to factors other than water: conductive minerals, graphite films on grain boundaries and enhanced intrinsic conductivity of mafic rocks have all been cited. Fluid immiscibility helps to overcome one theoretical obstacle to the hypothesis of deep crustal water, the need for an aquitard, but it does not explain the petrological and geochemical evidence which calls for dry conditions in stable lower crust. □

Bruce W. D. Yardley, of the Department of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK, is currently at the Department of Geology and Geophysics, University of Wisconsin, Madison, Wisconsin 53706, USA. John W. Valley is in the Department of Geology and Geophysics, University of Wisconsin, Madison, Wisconsin 53706, USA.

1. Bailey, R. C. *Nature* **371**, 238–240 (1994).
2. Bonham, L. C. *Am. Assoc. Petrol. Geol. Bull.* **64**, 549–567 (1980).
3. Green, A. G. et al. *Geophys. J.* **89**, 685–690 (1987).
4. Rumble, D. *Eur. J. Mineral.* **1**, 731–737 (1989).
5. Boden, A. & Eriksson, K. G. *Deep Drilling in Crystalline Bedrock*, Vol. 1 (Springer, Berlin, 1988).
6. Valley, J. W., Bohlen, S. R., Essene, E. J. & Lamb, W. J. *Petrol.* **31**, 555–596 (1990).
7. Morrison, J. & Valley, J. W. *J. Geol.* **99**, 559–570 (1991).
8. Banks, D. A., Davies, G. R., Yardley, B. W. D., McCraig, A. M. & Grant, N. T. *Geochim. cosmochim. Acta* **55**, 1021–1030 (1991).
9. Jenkin, G. R. T., Craw, D. & Fallick, A. E. *J. metamorph. Geol.* **12**, 429–444 (1994).

DAEDALUS

Corn circles

A TREE grows directly against the forces upon it. Thus the trunk climbs upwards against gravity, and the branches grow prestressed against gravitational sagging. Some palm trees show a spiral structure, which tends to be left-handed north of the Equator, and right-handed south of it. Atmospheric eddies show this same pattern of handedness, derived from the Earth's rotation. So Daedalus reckons that these trees are reacting, against the slight but persistent rotation applied to them by eddies in the local wind.

Daedalus hopes to test this notion by transplanting palms and other trees repeatedly between Britain and Australia, to see what they do. But he is also applying the torques directly. DREADCO foresters with ropes and windlasses are twisting the tops of selected trees, putting the whole trunk under a permanent torque. A tree grows by laying down a new outer layer of living wood over the existing core every year. Under torque, it should lay down a growing layer whose grain would not be vertical, but would climb helically upwards round the tree. Next year the foresters will reverse the torque: the next outer layer will form a reversed helix. By annually reversing a torque of the right value, successive helical layers of grain could be 'braided' at about 90°. This, of course, is the secret of plywood, whose successive plies are cross-grained in just this way. Daedalus has invented natural plywood.

Grain braiding should work on the branches as well as the trunk of the tree. DREADCO's natural plywood will be wonderfully strong in all directions, with no weak 'end grain'. When the humidity changes, it will swell and shrink evenly, without warping. Carpenters, yachtsmen and house builders will specify it for all their best work.

Agriculture could benefit too. Wheat, maize and rice, those vital cereals, have fragile stalks, easily broken by rough wind and weather. A braided helical fibre structure should strengthen them wonderfully. No farmer could apply individual reversing torques to millions of plants; but he could surround his field with vertical-axis wind turbines. Clockwise turbines would induce an anticlockwise eddy in the field. Reverse them, and they would induce a clockwise eddy instead. Repeat the process at intervals, and the growing crop would be strengthened by the helical braiding of the fibres of its stalks. It can now be revealed that the enigmatic corn circles in some British fields are the result of over-enthusiastic DREADCO pilot experiments.

David Jones

Palaeointensity puzzle

SIR — Valet and Meynadier¹ have published a palaeointensity record of the geomagnetic field for the past 4 Myr. This record shows substantial and apparently systematic structure between field reversals which, if real, has important implications for models of geomagnetic field generation.

Because the production rate of cosmogenic isotopes is, to a first approximation, inversely proportional to the square root of the geomagnetic field, profiles of these isotopes in geological reservoirs can also, in principle, be used to monitor palaeomagnetic field intensities². We demonstrated, for example, that the increase in ¹⁰Be content observed in a marine sediment during the Brunhes–Matuyama field reversal is consistent with a reduced field intensity during the reversal³. Others have used ¹⁰Be to support the argument that differences in ¹⁴C and calendar ages in pre-Holocene times are due to variations in the geomagnetic field⁴. If the intensity variations reported in ref. 1 are real (and if our understanding of the relationship between geomagnetic intensity and cosmogenic isotope production is sound), there should be observable changes in ¹⁰Be production.

Upper panel shows virtual axial dipole moments (VADM) derived by Valet and Meynadier¹ for samples studied here. Lower panel (dashed line) shows predicted ¹⁰Be/⁹Be values for samples studied here assuming palaeointensities of ref. 1 and relative production rates of Lal⁶ (see text). It should be noted that the calculations of Lal⁶ are for global production. The exact dependence of ¹⁰Be in a given sediment location depends on details of both atmospheric transport processes, and ¹⁰Be mixing processes in the ocean. These processes are insufficiently well understood at the moment to make quantitative corrections. However, because the sediment core studied is from the equatorial region, where geomagnetic influence on production is largest, it is likely that the predicted ¹⁰Be variations are actually underestimates of those that would occur at this site for the assumed palaeointensity variations. Data points represent ¹⁰Be/⁹Be measurements (corrected for radioactive decay of ¹⁰Be) for the samples studied here. In our previous work³, we measured ¹⁰Be in totally dissolved samples, and used ⁹Be in the same samples to try to normalize out possible variations caused by variable sediment composition and deposition rates. Here we have adopted a leaching technique, which ideally isolates ¹⁰Be and ⁹Be only from the 'authigenic' fraction of the sediment. We have shown (ref. 9, and our unpublished work) that this procedure produces more constant ¹⁰Be/⁹Be ratios in different sedimentary environments from the same time interval. We argue that, for the same reasons, this technique should also minimize non-production-related variations of ¹⁰Be/⁹Be over time at the same location. Samples consisting of ~1 g finely crushed sediment were agitated in 35 ml of 0.04 M hydroxylamine hydrochloride in 25% acetic acid at 96 °C for 6 hours⁹. After an aliquot was taken from the leachant for ⁹Be analyses, 0.5 mg ⁹Be carrier was added. Beryllium was extracted and purified using a series of acetyl acetone solvent extractions, then transformed into BeO by heating at 900 °C in a quartz crucible. ¹⁰Be/⁹Be measurements were carried out at the Gif-sur-Yvette accelerator mass spectrometry facility¹⁰, using NIST-SRM 4325 as a reference standard. ⁹Be measurements were made by flameless atomic absorption spectrometry (Perkin Elmer model 3030 with AS-60 automatic sample injector and Zeeman background correction) using standard additions, Mg(NO₃)₂ matrix modifier and platform atomization.

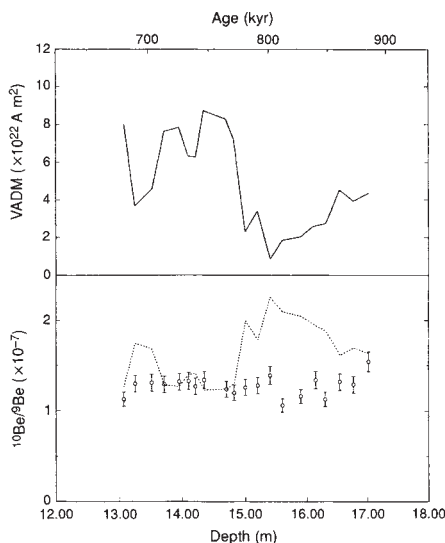
To test this hypothesis in the most direct possible way, we have measured the ¹⁰Be/⁹Be ratio in the same cubes from cores 851C and E analysed in ref. 1 (samples kindly provided by Valet and Meynadier). The time period chosen, about 650–900 kyr BP, brackets the most recent (Brunhes–Matuyama) field reversal, and corresponds to an interval in which Valet *et al.*⁵ found a similar variation of inferred palaeomagnetic intensity in several other marine cores.

Our results are shown in the figure as a function of equivalent depth in core 851D and estimated age of the sediment. Also shown are the palaeointensities deduced by Valet and Meynadier¹ for the same samples. In order to compare these with our results, it is necessary to calculate the expected effect of the inferred palaeointensity variations on ¹⁰Be production. To do this, we adopt the model calculations of Lal⁶, assuming that ¹⁰Be production will vary as his estimates of nuclear disintegration rates for an average solar modulation parameter $\phi = 450$. The calculations of O'Brien⁷ give quite similar results. The expected relative production rate of ¹⁰Be, normalized to the present-day virtual axial

dipole moment (8.2×10^{22} A m²), is shown as a dotted line.

As can be seen in the figure, the observed ¹⁰Be/⁹Be variations are far smaller than those expected on the basis of the palaeointensity estimates of ref. 1 for this period. A question that might be asked is why we do not see any increase in the ¹⁰Be/⁹Be ratio near the reversal itself, as seen in our earlier study³. Although samples were not selected with high time resolution in mind, a straightforward comparison with our earlier study would suggest that at least one or two points near the transition should show enhanced ¹⁰Be due to the low field associated with the reversal itself. Such a conclusion, however, rests on the assumption that bioturbation effects in the two cores have been identical. In fact, the core studied earlier (V16–58) showed varve-like structure near the reversal level, suggesting that bioturbation was relatively unimportant. By contrast, the present core shows a very homogeneous lithology, consistent with more extensive bioturbation. We are at present carrying out experiments on both cores to examine this question. Bioturbation should not have any influence on the problem under consideration here, however, as the time (and thus depth) scale of the variations being looked for are an order of magnitude longer than the reversal itself.

Are there any explanations for why the palaeointensities deduced from natural remanent magnetism (NRM)¹ might be erroneous? As discussed by Coe⁸, the hypothesis of partial overprinting following field reversals would offer a natural explanation for the sawtooth form of the changes observed. In such a scenario, when a field reversal occurs, a fraction of the magnetic grains previously deposited are imagined to be reoriented in the opposite direction, thus tending to diminish the net NRM signal recorded in the sediment. Because the fraction of grains reoriented decreases with depth, the net NRM signal would increase as one moves away from the reversal boundary, as observed. The main difficulty with such an explanation (ref. 8, and L. Meynadier *et al.*, manuscript in preparation) is that it requires some magnetic grains to be able



1. Valet, J. P. & Meynadier, L. *Nature* **366**, 234–238 (1993).
2. Raisbeck, G. M. & Yiou, F. *Nucl. Instrum. Meth.* **B5**, 91–99 (1984).
3. Raisbeck, G. M., Yiou, F., Bourles, D. & Kent, D. V. *Nature* **315**, 315–317 (1985).
4. Lao, Y. *et al.* *Nature* **357**, 576–578 (1992).
5. Valet, J. P., Meynadier, L., Bassinat, J. C. & Garnier, F. *Geophys. Res. Lett.* **21**, 485–488 (1994).
6. Lal, D. in *Solar–Terrestrial Relationships and the Earth Environment in the Last Millennium* (ed. Cini Castagnoli, G.) 216–241 (North Holland, 1988).
7. O'Brien, K. J. *geophys. Res.* **84**, 423–431 (1979).
8. Coe, R. *Nature* **366**, 205–206 (1993).
9. Bourles, D., Raisbeck, G. M. & Yiou, F. *Geochim. cosmochim. Acta* **53**, 443–452 (1989).
10. Raisbeck, G. M., Yiou, F., Bourles, D., Lestranguez, J. & Deboffe, D. *Nucl. Instrum. Meth.* **B29**, 22–26 (1987).

to reorient at depths of up to several metres, which is at least an order of magnitude larger than the depth previously accepted for the 'locking-in' of NRM.

Thus, in at least one marine core the intensities of the palaeomagnetic field inferred from NRM and ^{10}Be are incompatible. Further investigation will be necessary to decide which parameter is giving misleading results, and why.

G. M. Ralsbeck

F. Ylou

S. Z. Zhou

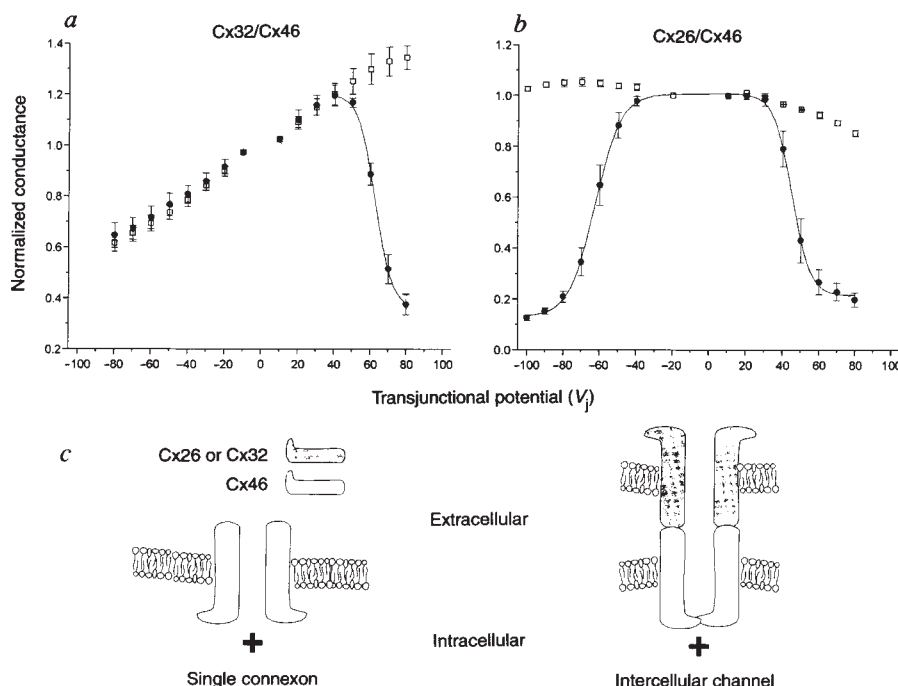
Centre de Spectrométrie Nucléaire et de Spectrométrie de Masse,
IN2P3-CNRS Bât. 108,
91405 Orsay Campus, France

Voltage gating of connexins

SIR — Connexins (Cx) oligomerize into channels (called connexons) that span a single plasma membrane. Connexons in adjacent cells align to form complete intercellular channels that are sensitive to the voltage difference between the cytoplasm of the coupled cells. Because the intercellular channel spans two plasma membranes, adjacent cells can contribute different connexins to form a heterotypic intercellular channel. It is not known if the polarity of voltage gating of an intercellular channel is determined by an intrinsic voltage sensitivity of individual connexons.

Early models described the voltage dependence of intercellular channels as arising from two independent gates, one contributed by each connexon, that would close for a transjunctional voltage (V_j) of a given polarity¹. The apparent deviation of heterotypic Cx26/Cx32 channels from this rule² has been explained by the recent demonstration that Cx26 and Cx32 have opposite polarities of voltage gating in intercellular channels^{3,4}. Cx26 connexons close when their cytoplasmic end is relatively positive and open for relative negativity, whereas Cx32 connexons are gated in an opposite fashion. Pairing Cx26 with Cx32 in a heterotypic channel results in steady-state rectification², due to these opposite polarities of connexon closure.

It is possible to test directly whether the two voltage gates of an intercellular channel operate independently with respect to polarity by examining the physiological properties of heterotypic channels containing Cx46. The voltage-gating polarity of Cx46 is unambiguously defined because it forms functional connexons that can be opened by depolarization in single membranes^{5,6}. Intercellular channels containing only Cx46 are symmetrically gated



Connexons composed of Cx46 close for relative positivity in heterotypic intercellular channels. Initial ($\square$) and steady-state ($\bullet$) conductances were plotted at different transjunctional potentials for heterotypic channels composed of Cx32, Cx46 and Cx26. Positive and negative V_j values correspond to depolarization and hyperpolarization, respectively, of the cell injected with the connexin written on the right, that is, Cx46. a, Cx32/Cx46 heterotypic channels exhibit notable instantaneous rectification, and a loss of steady-state sensitivity when the Cx32 expressing cell is relatively positive. b, Cx26/Cx46 heterotypic channels show steady-state voltage dependence for both polarities of V_j . Solid lines represent fitted Boltzmann equations. Results are shown as means $\pm$ s.e.m. c, Schematic representation of voltage gating polarity of Cx46 in single connexons and intercellular channels. In single *Xenopus* oocytes, Cx46 forms voltage gated connexons which open on depolarization and close on hyperpolarization of the cell. In this case, single Cx46 connexons open for relative positivity at the cytoplasmic end of the channel. In heterotypic channels with either Cx26 or Cx32, the Cx46 connexons close on depolarization of the Cx46-expressing cell, and are open for hyperpolarization of the Cx46-expressing cell. When incorporated into intercellular channels, Cx46 connexons reverse their polarity of voltage gating, and close for relative positivity at the cytoplasmic end of the channel.

by voltage⁷, which does not allow the determination of polarity in individual connexons containing Cx46. We have used the paired-oocyte expression system to examine the interaction of Cx46 with either Cx32 or Cx26, whose polarity of voltage gating has been inferred from studies of intercellular channels^{3,4}.

HETEROTYPIC CHANNEL FORMATION BY Cx26		
Pairing (cell 1/cell 2)	Junctional conductance (μS) (mean $\pm$ s.e.m.)	No. of pairs
Cx32/Cx46	3.41 ± 0.75	8
Cx26/Cx46	5.87 ± 1.53	9

Cx46 formed heterotypic channels with both Cx32 and Cx26 (see table), which allowed the direct comparison of the polarity of voltage gating of Cx46 in intercellular channels with Cx46 connexons in single membranes. In heterotypic Cx32/Cx46 intercellular channels, Cx46 had the same apparent polarity of gating as Cx26 in heterotypic Cx32/Cx26 channels, closing for relative positivity (a in the figure). In addition, Cx26/Cx46 heterotypic channels showed reduced steady-state conductance for both positive and negative transjunctional potentials (b). Thus, Cx46 connexons in heterotypic Cx26/Cx46 channels must also close for relative posi-

tivity at their cytoplasmic end to produce the observed closure for both polarities of V_j . In both cases, Cx46 connexons in intercellular channels are closing for relative positivity at their cytoplasmic end, an opposite polarity of closure to that of single Cx46 connexons. The polarity of voltage gating for single Cx46 connexons and Cx46 in intercellular channels is summarized in c in the figure. The simplest explanation is that the polarity of voltage gating for Cx46 connexons reverses when they become incorporated into an intercellular channel.

Theoretically, single connexons could generate functional diversity by containing more than one type of connexin. However, the structural nature of intercellular channels offers an alternative mechanism to generate functional diversi-

- Harris, A. L., Spray, D. C. & Bennett, M. V. L. *J. gen. Physiol.* **77**, 95–117 (1981).
- Barrio, L. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8410–8414 (1991).
- Bennett, M. V. L., Rubin, J. B., Bargiello, T. A. & Verselis, V. K. *J. Physiol.* **43**(suppl.), S301–S310 (1993).
- Verselis, V. K., Ginter, C. S. & Bargiello, T. A. *Nature* **368**, 348–351 (1994).
- Paul, D. L., Ebihara, L., Takemoto, L. J., Swenson, K. I. & Goodenough, D. A. *J. Cell Biol.* **115**, 1077–1089 (1991).
- Ebihara, L. & Steiner, E. *J. gen. Physiol.* **102**, 59–74 (1993).
- White, T. W., Bruzzone, R., Wolfram, S., Paul, D. L. & Goodenough, D. A. *J. Cell Biol.* **125**, 879–892 (1994).

ty, in that allosteric interactions between opposing connexons could generate novel properties. The polarity of voltage gating of Cx46-containing connexons is reversed upon incorporation into an intercellular channel. These results contradict the theory of independent voltage gates and provide a basis for a new model of voltage gating in gap junctions where the voltage sensor of an intercellular channel is not intrinsic to a single connexon, but is shared between, and modulated by, the composition of opposing connexons.

Thomas W. White

Roberto Bruzzone

Daniel A. Goodenough

Department of Cell Biology,

David L. Paul

Department of Neurobiology,

Harvard Medical School,

Boston, Massachusetts 02115, USA

Extraordinary salmon growth

SIR — Over the past decade, transgenic technologies have been explored in various fish species¹, with particular emphasis on growth enhancement as a strategy to shorten long production cycles. The original gene construct used to double the size of mice used a mouse metallothionein promoter controlling a rat growth hormone gene². This and other mammalian constructs have had no or modest effects on growth in transgenic fish^{3–5}, prompting the development of fish gene constructs^{6,7} with improved effects on growth performance⁷. In addition, public concern over the use of DNA from non-homologous sources makes it desirable to develop constructs from as close to homologous DNA as is practical.

For the production of transgenic Pacific

salmon, we have circumvented some of these problems by developing a gene construct (pOnMTGH1) where all genetic elements were derived from sockeye salmon. This 'all-salmon' construct consists of the metallothionein-B promoter⁸ fused to the full-length type-1 growth hormone gene⁹, and is identical in concept to the construct first used to stimulate growth in transgenic mice².

We microinjected linear pOnMTGH1 DNA into the blastodisc region (animal pole) of coho salmon eggs that were developmentally arrested immediately after fertilization. From more than 3,000 eggs injected, we found that 6.2% of the individuals surviving to one year of age retained pOnMTGH1 DNA in their fin tissue. Control uninjected salmon displayed a uniform frequency distribution of weight classes (Fig. 1a). The group microinjected with pOnMTGH1 had the same modal weight as controls, but, in addition, contained many larger individuals that clearly lay outside the normal distribution (Fig. 1b). Most large individuals were transgenic in fin tissue, indicating that the presence of the pOnMTGH1 construct was responsible for the growth enhancement. Of the large fish that did not have pOnMTGH1 DNA in their fin tissue, 54 per cent were found to be positive in blood cells, confirming the mosaic nature of first-generation transgenic fish. On average, the trans-

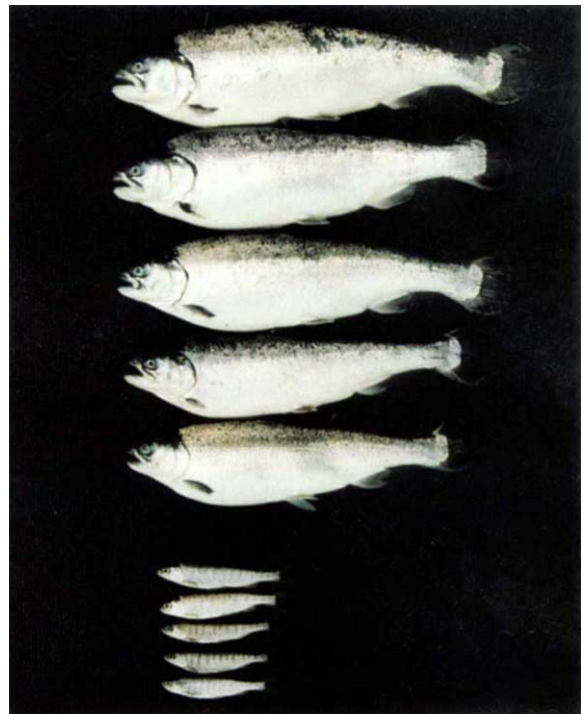


FIG. 2 Non-transgenic (bottom) and transgenic (top) coho salmon siblings at 14 months of age showing size difference and silver appearance of transgenic individuals indicative of transformation to seawater adaptability. Length of top large fish (fork length), 41.8 cm.

genic salmon were more than 11-fold heavier than non-transgenic controls, with a range from no growth stimulation to one individual 37 times larger than controls.

In addition to growth acceleration, most transgenic fish precociously developed a silver body coloration (Fig. 2) typical of salmon undergoing the physiological pre-adaptation (smoltification) necessary for the spring migration from fresh water to the marine environment. Levels of serum growth hormone in juvenile coho salmon are normally very low, then rise in the spring in correlation with smoltification¹⁰. In our experiments, winter levels of growth hormone in transgenic fish were more than 40-fold higher than in non-transgenic controls, indicating that 'unregulated' overexpression of growth hormone from pOnMTGH1 and its consequent physiological effects had

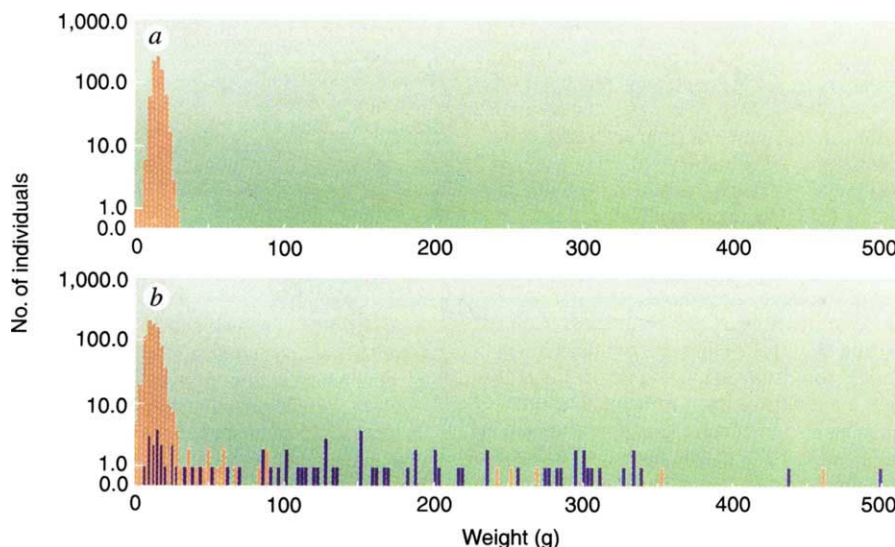


FIG. 1 Frequency distributions of coho salmon weights at 12 months post-fertilization: *a*, control ($n = 792$); *b*, microinjected with pOnMTGH1 ($n = 1,073$). Transgenic individuals identified by PCR are indicated by purple bars in *b*.

1. Fletcher, G. & Davies, P. L. *Genet. Engng* **13**, 331–369 (1991).
2. Palmiter, R. D. *et al.* *Nature* **294**, 611–615 (1982).
3. Guyomard, R., Chourrou, D., Leroux, C., Houdebine, L.-M. & Pourrain, F. *Biochimie* **71**, 857–863 (1989).
4. Zhu, Z. in *Transgenic Fish* (eds Hew, C. & Fletcher, G.) 92–119 (World Scientific, Singapore, 1992).
5. Lu, J.-K., Chen, T. T., Chrisman, C. L., Andrisini, O. M. & Dixon, J. E. *Molec. mar. Biol. Biotech.* **1**, 366–375 (1992).
6. Cavari, B., Hong, Y., Funkenstein, B., Moav, B. & Schartl, M. *Fish Physiol. Biochem.* **11**, 345–352 (1993).
7. Du, S. J. *et al.* *Bio/Technology* **10**, 176–180 (1992).
8. Chan, W. K. & Devlin, R. H. *Molec. mar. Biol. Biotech.* **2**, 308–318 (1994).
9. Devlin, R. H. *Can. J. Fish. aquat. Sci.* **50**, 1738–1748 (1993).
10. Sweeting, R. M., Wagner, G. F. & McKeown, B. A. *Aquaculture* **45**, 185–197 (1985).

been seasonally uncoupled from normal pituitary-derived production.

Although we are encouraged by the growth results observed with Pacific salmon, it remains to be seen whether pOnMTGH1 and other piscine gene constructs will generally function better than non-homologous ones in other cold- and warm-water fish species where growth hormone is limiting with regard to growth performance.

Robert H. Devlin*, Timothy Y. Yesaki,
Carlo A. Blagi, Edward M. Donaldson

Fisheries and Oceans Canada,
4160 Marine Drive,
West Vancouver,
British Columbia V7V 1N6, Canada

Penny Swanson

Northwest Fisheries Science Center,
National Marine Fisheries Service,
Seattle, Washington 98112-2097, USA

Woon-Khiong Chan

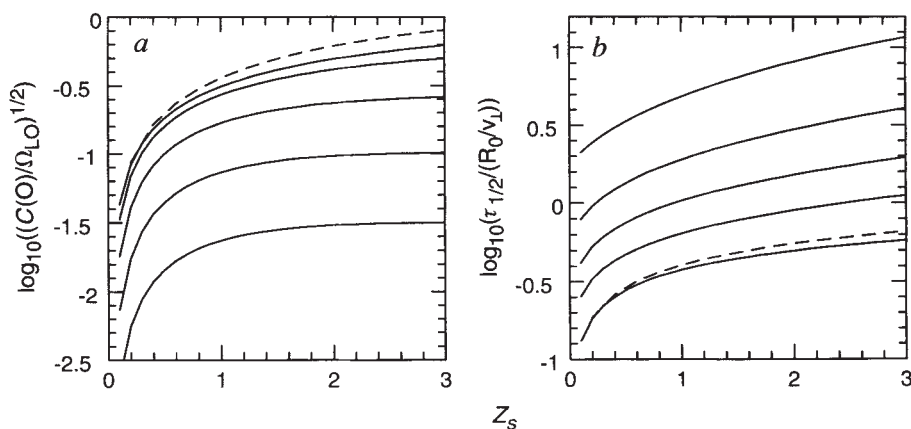
Institute of Molecular and Cell Biology
and Department of Zoology,
National University of Singapore,
Singapore 0511

* To whom correspondence should be addressed.

Quasar variability

SIR—Hawkins¹ monitored the observed variability of quasars with redshifts between 1 and 3, and finds that the timescale of variability decreases with redshift, which he argues is evidence in favour of the quasar light being microlensed by sub-stellar objects along the line of sight. I show here that in fact the timescale of lensing should increase with redshift, because the average redshift of the lensing objects increases with the redshift of the quasar, resulting in a longer variability timescale as the time taken for the object to cross the line of sight is increased by time dilation.

I have calculated the autocorrelation function of the quasar magnitude variations produced by gravitational lensing by a population of intervening, randomly distributed, compact objects, under the assumption that the amplifications A due to lensing by different objects combine multiplicatively. This is exact in the limit of very small lensing probability, and has been shown to be a good approximation up to modest lensing optical depths². It is expected to be a reasonable approximation in the present case. The magnitude change is $\mu \equiv -2.5 \log_{10} A$, so that the contributions of different objects to μ add. For the case that the quasar is a point source at redshift z_s , lensing by a point mass M_L at redshift z_L and distance l from the line of sight amplifies the brightness by a factor $A(u) = (u^2 + 2)/u\sqrt{u^2 + 4}$, where $u = l/R_E$, and $R_E(M_L, z_L, z_s)$ is the Einstein ring radius³. The amplification varies with time due to motions of the observer, source and lens; here I consider only the latter, with lenses moving with a



a, Root-mean-square magnitude variation, $C^{1/2}(0)$; b, variability timescale, $\tau_{1/2}$, as functions of z_s , for different source radii R_s . Solid curves are for $\Omega_0=1$ and $R_s/R_0=0, 0.3, 1, 3, 10$ (decreasing vertically for $C(0)$ and increasing vertically for $\tau_{1/2}$). Dashed curves are for $\Omega_0=0.1$ and $R_s/R_0=0$.

transverse velocity $v_\perp$. The autocorrelation function at time lag τ for a single lens is $\int_{-\infty}^{\infty} \mu(t+\tau)\mu(t) dt$. This is then integrated over l and z_L , assuming a constant comoving number density of lenses, to give the net autocorrelation function $C(\tau) \equiv \langle (\mu(t+\tau) - \langle \mu \rangle)(\mu(t) - \langle \mu \rangle) \rangle$. The characteristic timescale of magnitude variations due to lenses at redshift z_L is just $(1+z_L)R_E/v_\perp$, where the factor $(1+z_L)$ results from time dilation.

The autocorrelation function $C(\tau)$, which is symmetrically peaked around $\tau=0$, can be simply characterized by an amplitude and a timescale. The amplitude $C(0) = \langle (\mu - \langle \mu \rangle)^2 \rangle = 0.806 p_L$ gives the mean square magnitude variation. Here, $p_L(z_s, \Omega_{L0})$ is the probability for a strong lensing event ($l < R_E$), also called the optical depth for lensing³, and is proportional to the fraction of the critical density in lensing objects, $\Omega_{L0} = 8\pi G n_{L0} M_L / (3H_0^2)$, but independent of the lens mass M_L . (n_{L0} is the present number density of lenses, and H_0 is the present Hubble constant, taken to be $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$). The timescale, for example the half-width at half maximum $\tau_{1/2}$, is independent of Ω_{L0} , but scales as $\tau_{1/2} \propto R_0/v_\perp$ for given z_s . Here, $R_0 \equiv (4GM_L/cH_0)^{1/2}$ is the characteristic value of R_E , and the characteristic timescale $R_0/v_\perp = 2.6 \text{ yr} (M_L/10^{-3} M_\odot)^{1/2} / (v_\perp/400 \text{ km s}^{-1})$. The figure shows the dependence of $C^{1/2}(0)$ and $\tau_{1/2}$ on z_s , for total density parameters $\Omega_0=0.1$ and $\Omega_0=1$.

I have also computed the effect of a finite source size, approximating the quasar as a uniform surface brightness disk of radius R_s , following P. Schneider (personal communication). The radii of the optical continuum emitting regions of quasars are uncertain, but are thought to be at least 10^{15} cm for the more luminous quasars. The results are shown in the figure for different values of $R_s/R_0=0.3(R_s/10^{15} \text{ cm})/(M_L/10^{-3} M_\odot)^{1/2}$. The effect of a finite source size is to decrease the amplitude of magnitude variations and increase their timescale. For $R_s/R_0 \gg 1$, these scale

as $C^{1/2}(0) \propto R_s^{-1}$ and $\tau_{1/2} \propto R_s$; only a fraction of the quasar light is significantly amplified, and the timescale is essentially that for a lensing object to cross the quasar disk.

Hawkins estimates a quantity similar to $\tau_{1/2}$ from his observational data, and finds $\tau_{1/2} \sim 1\text{--}5 \text{ yr}$, with the timescale decreasing by a factor 0.6–0.7 from $z_s \approx 1.5$ to 2.5. This is in conflict with the gravitational lensing calculations, which predict that $\tau_{1/2}$ should increase by a factor ranging from 1.2 for $R_s/R_0 \ll 1$ to 1.4–1.6 for $R_s/R_0 \gg 1$, for Ω_0 in the range 0.1–1. The simplest intrinsic variability model predicts that $\tau_{1/2}$ should increase as $(1+z_s)$, that is, a factor 1.4, due to time dilation at the source. In this sense, the lensing model (for point sources) is closer to the observed trend. Another test of the lensing hypothesis would be to compare predicted and observed autocorrelation amplitudes $C(0)$. It would also be useful to carry out numerical simulations of lensing, including the various observational selection effects, to check the accuracy of the assumptions I have made here.

C. G. Lacey

Physics Department,
University of Oxford,
Oxford OX1 3RH, UK

1. Hawkins, M. R. S. *Nature* **366**, 242–245 (1993).
2. Pei, Y. C. *Astrophys. J.* **403**, 7–19 (1993).
3. Paczynski, B. *Astrophys. J.* **304**, L1–L5 (1986).

Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.

Presence of our past

Jared M. Diamond

The History and Geography of Human Genes. By Luigi Luca Cavalli-Sforza, Paolo Menozzi and Alberto Piazza. Princeton University Press: 1994. Pp. 541. \$150, £150.

GREAT science often depends on a mixture of inspiration and perspiration. As an example, in *Nature* 329, 741 (1987), Alberto Piazza and colleagues reported a novel approach to the difficult basic problem of estimating migration rates among human populations. Their inspired solution was to look for evidence of migration in the geographical distribution of surnames, much as one would study the distribution of neutral alleles of a genetic locus. But perspiration was also necessary: the authors analysed the complete telephone directories of 91 Italian provinces, after eliminating names of commercial subscribers. The result, validated by comparison with migration rates calculated directly from Italian records of registration for residency: estimates of male immigration rates based on an enormous sample (more than 10,000,000 Italian telephone owners).

Two of these inspired and perspiring geneticists, Luigi Luca Cavalli-Sforza and Alberto Piazza, have now joined with Paolo Menozzi in an even more monumental study of human population genetics. The book is in part a synthesis, by far the most complete one ever attempted, of current knowledge of geographical variation in human gene frequencies. Two tabular appendices summarize decades of measurements of 128 alleles in 491 human populations and 116 population aggregates, while 518 maps visually convey the distributions of individual alleles and 8 colour maps synthesize gene-frequency gradients for each continent. Five-hundred-and-forty-one pages of text analyse how these distributions and gradients may have resulted from migration, admixture, drift and selection.

Other disciplines besides genetics provide information useful for deducing human migrations, however. Correspondingly, genetic information can help answer important questions in other disciplines. Thus the book was in fact written to be an encyclopaedia not only of world genetics but also of world archaeology, geography, history and prehistory, linguistics and physical anthropology. Five of the chapters synthesize, for each continent in turn, these disparate sources of information. These other bodies of information are not added as mere window-dressing to an ordinary text of genetics. For instance, the chapter on Asia starts with a couple of pages on Asian geography, goes on to summarize in 22 pages the prehistory and history of each of six parts

of the Asian continent, and devotes 3 pages each to Asian physical anthropology and Asian linguistics before culminating in a 30-page interpretation of Asian genetics.

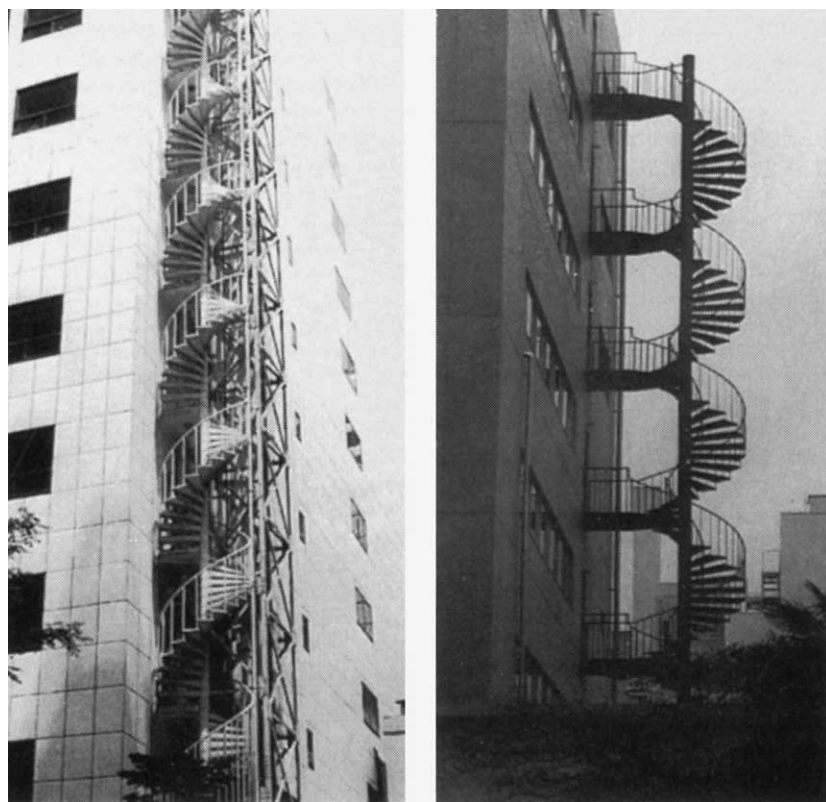
In reading this book, I felt as I do when eating one of my wife's raspberry chocolate nut layer cakes. Both the cake and the book are rich, dense, uniformly high in quality, heterogeneous in subject matter and best enjoyed by repeatedly removing from the refrigerator or bookshelf to take small bites. The authors' accounts of two populations, the Ainu and the Sardinians, will give prospective readers a sense of the book's flavour.

Perhaps more anthropological attention has been devoted to the Ainu of northern Japan than to any other aboriginal people in the world. They have drawn this interest especially because they look more 'Caucasoid' than Japanese, being very hairy, lacking epicanthic eye folds and often having European-like hair colour, skin colour and hair form. De-

pending on which linguist one believes, their distinctive language is either of indeterminate affinities or is a very isolated member of the Altaic language family.

Japan's archaeological history includes the world's most ancient pottery and a long Neolithic period ending with the arrival of rice and iron from the Asian mainland around 400 BC. This history is sometimes interpreted in terms of original Ainu inhabitants being displaced only recently by the modern Japanese. Various complex schemes have been proposed to explain how ancient Caucasoids might have ended up in Japan to become the Ainu. Cavalli-Sforza and colleagues ably summarize all of these strands, then show that in their genes the Ainu are northeast Asians most closely related to the Japanese, Koreans and Tungus. As the authors note: "It seems reasonable to discard the myth of a Caucasoid origin of the Ainu". The Ainu's renowned hairiness and other distinctions of appearance are evidently only skin-deep and may have evolved through sexual selection during long isolation on the Japanese archipelago.

As for Sardinia, its human settlement dates back to the Late Palaeolithic at least 9,000 years ago. In the past 3,000 years the island has been conquered by successive waves of Phoenicians, Romans, Vandals, Byzantines, Arabs, Spaniards and Ita-



SPIRAL staircases in Seoul, Korea (left), and Kyoto, Japan (right). From *Symmetry: A Unifying Concept* by I. and M. Hargittai. It contains 550 photos and 300 drawings collected by these two structural chemists. Shelter, \$18 (pbk).

lians. Sardinia's language today is a distinctive member of the Romance subgroup of the Indo-European language family, but some words suggest pre-Indo-European legacies. Genetically, modern Sardinians are the most divergent European population after the Lapps, exceeding even the Basques in their distinctiveness. They have the world's highest frequency of the blood-group allele MNS*M (78 per cent), a thalassaemia variant rare elsewhere, a peculiar HLA pattern and the Mediterranean's lowest frequency of rhesus-negative genes.

Cavalli-Sforza and colleagues argue that these distinctions stem from thousands of years of genetic drift in a small population, as Sardinia is the most remote main Mediterranean island and its pre-Neolithic population could have numbered only about a thousand. The many invasions of the past few millennia have stamped mainland political control and a

new language on Sardinia but have left only limited genetic traces.

These two nuggets must suffice to hint at the richness of this raspberry chocolate cake of human sciences. Prospective buyers may wince at the price, until they see what they are getting for it. The book weighs seven pounds, contains more than 1,000 pages in a large format, and is much cheaper than the summed cost of all the other books and photocopied articles that it supercedes. It represents a unique database and unique synthesis. No geneticist, evolutionary biologist, linguist, archaeologist or prehistorian can afford not to own it, except perhaps those with very narrow interests. I own two copies, one kept in my office and the other at home. □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024-1751, USA.

A brief history of time warps

Bernard Carr

Black Holes and Time Warps: Einstein's Outrageous Legacy. By Kip. S. Thorne. Norton/Picador: 1994. Pp. 600. \$30, £20.

ALTHOUGH the theory of general relativity is nearly 80 years old, only in the past 30 years has evidence mounted that its more peculiar predictions are relevant to the real world. The discovery of pulsars and quasars and other exotic astronomical sources has demonstrated that Einstein's theory is more than a mathematical curiosity — it is a vital ingredient of any attempt to understand the Universe. Nowhere is this more true than in the context of Einstein's most outrageous legacy, black holes. These strange objects are now invoked everywhere: as the ghosts of massive stars, as the powerhouse of active galactic nuclei, even perhaps as the gateways to other universes — so it is no surprise that they have been the focus of so many popular presentations.

Just as intriguing as the evidence for black holes and their peculiar properties, however, is the history of the subject, and this neglected but engrossing area is the prime focus of this book. Indeed, the buzzwords of the title rather understate its scope. For it is not just a book about black holes *per se*; it is also about the people who pioneered black-hole research and about the sociological and political factors that influenced the development of their ideas. As such, it is a scholarly and enthralling work, written by somebody who can offer a unique perspective. For Thorne has devoted most of his life to studying black holes and knows most of the key charac-

ters. Indeed, much of the material is based on extensive interviews with them. Although it is non-technical, at 600 pages the book may be rather daunting for the average layperson. But it also has some important messages about the nature of science in general, so it is not only black-hole specialists who will be stimulated.

A recurring theme is the apparent outrageousness of some of the predictions of general relativity, often resulting in resistance from leading physicists of the day. Thus we learn of Einstein's resistance to black holes, of Eddington's resistance to Chandrasekhar's white-dwarf calculations, of Oppenheimer's resistance to Zwicky's speculations about neutron stars and of Wheeler's initial resistance to Oppenheimer's picture of gravitational collapse. Perhaps the most outrageous prediction of all is that wormholes (not the same as black holes but closely related) can act as time-machines, a notion that Thorne himself helped to pioneer and which led a friend of his to worry that he was cracking up! Indeed, perhaps a certain sense of defensiveness underlies Thorne's emphasis on the frequent scepticism of establishment physics about relativity's predictions.

It is interesting that what makes outrageous ideas acceptable — even before unambiguous evidence has accumulated to support them — is the backing of a few influential people. Thus Zwicky's speculations about neutron stars were taken seriously only after the Russian physicist Landau wrote about them (although Landau's model turned out to be wrong) and speculations about time-machines became respectable only after a paper by

Thorne and colleagues "broke the barrier". (Had the paper been written by a less renowned figure, it would probably have had a rougher ride.) Truth may win out in the end but there is no doubt that how long it takes to do so depends on the attitudes of a small number of influential people who guide (and sometimes misguide) the direction of mainstream research.

Another recurring theme is how often great people are wrong — at least initially. Thus Hawking dismissed Bekenstein's ideas of black-hole entropy, Ze'ldovich dismissed Hawking's theory of black-hole radiation, and Khalatnikov and colleagues dismissed Penrose and Hawking's claim that the Universe must have started with a singularity. Being brilliant is clearly no guarantee of being correct, as in all these examples the dismissal turned out to be unfounded. Perhaps the real hallmark of great scientists is that they are prepared to admit their mistakes and maybe that is why Thorne is happy to admit so many of his own: his resistance to Wheeler's idea of matter being turned into radiation at a singularity, his disagreement with Ze'ldovich about radiation from spinning black holes and his slowness in seeing the significance of Braginsky's ideas on the quantum limits of gravitational-wave detectors. Of course, contention is a necessary part of science — it provides a vital spark in the drive towards understanding. Sometimes disagreements are intense but more often they are good-natured and humorous, as illustrated by the numerous bets that relativists and astrophysicists seem to indulge in.

Thorne also emphasizes the connection between science and politics. For example, we learn how Ginzburg worked on cosmic rays because his wife was falsely accused of trying to kill Stalin and how Landau was motivated to publicize his thoughts on neutron stars because he hoped the splash of publicity would protect him from arrest and death. Ironically, international tensions are often a spur to the progress of science and many astronomical advances have been aided by military pressures (radioastronomy by the development of radar, gamma-ray astronomy by the need to detect terrestrial nuclear explosions). Indeed the recent cuts in US science spending emphasize how much science has suffered since the end of the Cold War. Sometimes military distractions can be beneficial in unexpected ways; although work on the problem of stellar collapse was interrupted by the Second World War, the involvement of people like Ze'ldovich and Wheeler in nuclear-weapon programmes ultimately speeded its solution because the bomb codes could be adapted to stellar-implosion calculations.

Another feature of scientific progress that is clearly illustrated by the black-hole

story is how much it depends on the people involved and their relationships with each other. Of all such relationships, none is more crucial than that between student and mentor. Science progresses rather like evolution: at each stage, key individuals train the next generation of researchers who then perpetuate the ideas and the research style that constitute their intellectual genes. Every researcher plays a part in this process but some have been particularly influential, and Thorne justly highlights the contributions and different styles of three outstanding mentors: Wheeler (the inspirational visionary), Ze'ldovich (the hard-driving coach) and Sciama (the self-sacrificing catalyst).

Focusing on the people involved in the black-hole story is illuminating because it sheds light on the quirky circumstances that have sometimes triggered important discoveries. Thus we learn how Hawking's black-hole area theorem came to him while getting into bed, how Penrose's singularity theorem struck him while crossing the road and how he got interested in cosmology because his brother shared an office with Sciama, how the vital Soviet collaboration between Novikov and Ze'ldovich was triggered by a chance remark of Novikov's wife, and how a science-fiction novel by Sagan inspired Thorne's work on time-machines. One often imagines that the frontiers of science advance systematically, powered by the inexorable drive of reason and new data, but it is surprising how often crucial advances seem to hinge on serendipity and chance encounters.

One could level a few criticisms at the book. Its tone is lowered somewhat by the prologue: a science-fiction story involving a futuristic tour of black holes. This is rather out of keeping with the scholarly style of the rest of the book and seems to have been included to whet the appetite of the layperson. (It also seems rather implausible that androids in the next century could face destruction as a result of forgetting physics familiar to every student of relativity in this century!) The book is heavily autobiographical in parts (especially towards the end), with considerable emphasis placed on the work of the author and his colleagues. There is no denying the important contribution of Thorne's group but there is inevitably a certain subjectivity in selecting the highlights of a subject and several workers will be disappointed to find their contributions to the field omitted. Nevertheless, for the most part the book is a masterly, well-balanced and comprehensive overview of one of the most exciting fields of twentieth-century physics. □

Bernard Carr is in the Astronomy Unit, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, UK.



An adolescent cheetah, showing the long light hairs on her nape. The social behaviour of these endangered carnivores is complex: cubs stay with their mother for a full year after weaning; adolescents remain in groups; adult males live in permanent associations with each other; and adult females live alone. In *Cheetahs of the Serengeti Plains: Group Living in an Asocial Species*, T.M. Caro collates more than a decade of research in the wild, providing a detailed account of the behaviour and ecology of the cheetah. University of Chicago Press, \$70, £55.95 (hbk), \$26.95, £21.50 (pbk).

Life's complexity

Stuart A. Newman

How the Leopard Changed Its Spots: The Evolution of Complexity. By Brian Goodwin. Weidenfeld and Nicolson/Scribners: 1994. Pp. 233. £18.99, \$23. To be published in the US in October.

NATURE, long stigmatized as being red in tooth and claw, is actually Green, according to Brian Goodwin's new book. The leopard of the title is the science of biology, which the author asserts is undergoing a major change in theoretical perspective, one in which organisms will "cease to be simply survival machines and assume intrinsic value, having worth in and of themselves, like works of art." It is difficult to take exception to the book's point of departure, that the reductionist paradigm in organismal biology, for all its successes, is unable to provide a compelling account of the evolutionary origins of form and function. It is equally hard to disagree with its conclusion that the free reign of capital, stultifying political institutions and a "monoculture mentality" that emphasizes quantity over diversity

are destroying the natural world and human prospects. But Goodwin's aim is to establish connections between these realms and to propose a common basis for their renewal. This is a large order, in which he is only partly successful.

Goodwin is preeminent among critics of the gene-centred biological science that has flourished over the past 40 years. He decries the tunnel vision that attributes all important organismic properties to expression of genetic content. Making use of analogies to physics and chemistry, he shows the inadequacy of the proposition that composition uniquely determines form. A collection of carbon atoms, for example, might constitute a diamond, a lump of graphite or a fullerene. A volume of liquid water can be stationary or form waves or vortices. Similarly, the stereotypical arrangements of leaves on a plant stem may represent the variant results of a common physical process. The external world leaves an imprint on all matter, nonliving and living, so that the forms that occur are outcomes of an interplay between intrinsic and extrinsic properties.

Goodwin rejects accounts of biological organization that rely principally on history (such as Darwin's notion that inheri-

tance from a common ancestor is explanatory of form). Much of his presentation here draws heavily on the language of dynamical systems theory and the "new sciences of complexity" associated most closely with the Santa Fe Institute in New Mexico.

Complexity theory looks for commonalities among the relatively abstract systems susceptible to computer modelling, seeking 'generic forms' that may emerge out of synchronic brews of interacting components. One problem with this view is that real physical and biological systems are made of distinct kinds of materials. If assemblages of electrons, protons and neutrons, or liver cells or ants, have any generic forms in common, they are unlikely to be the most significant properties of these systems. In several of the examples discussed (for example, a mechanochemical model of vertebrate limb development that was experimentally disconfirmed several years ago) there is not enough attention given to the specificity of the materials under consideration.

Goodwin concedes that living systems are distinguished from nonliving systems, no matter how complex, by the presence of "powerful particulars that give them the capacity to regenerate and reproduce their own natures under particular conditions". So we are brought back, despite the author's intentions, to genes and history as the distinguishing characteristics of organisms. Whatever the evolutionary origin may have been of particular organismal forms (and the dynamical processes discussed here are a good bet for many of them), the present-day developmental realization of these forms must depend greatly on an accumulation of nongeneric molecular circuitry.

As Goodwin turns his attention from developmental biology to animal and human behaviour and to social ecology, he proposes a "science of qualities", as distinct from, and complementary to, the "science of quantities", which, he suggests, has prevailed since Galileo. Qualities in his terminology are the "expression of an integrated whole", but they then become conflated with quality as value, as in "quality of life". This is particularly evident in the contrast he draws between the healthy and environmentally correct Hunza people of Himalayan Pakistan and some messed-over populations in technologically advanced Western societies. While Goodwin makes some telling points about the irrelevance of genetics in accounting for cultural difference, societies, like organisms, are laden with their various histories. Concerning the means of liberation from such constraints, he has little to say. □

Stuart A. Newman is in the Department of Cell Biology and Anatomy, New York Medical College, Valhalla, New York 10595, USA.

Myths dispelled

John Mann

The Consumer's Good Chemical Guide.

By John Emsley. W. H. Freeman/Spektrum: 1994. Pp. 374. £18.99, \$24.95. To be published in the US on 27 October.

CHEMISTRY and chemicals receive a bad press. The very word 'chemical' now has a certain malevolence, conjuring up fears about pesticide residues in food, pollution of rivers and the atmosphere, or harmful drugs. Chemists must indeed take the blame for the invention of tetraethyllead, chlorofluorocarbons and nerve gases, to name but a few. But lest we forget the positive contributions of chemistry, it is worth noting that more than 20 million people died from bacterial infections contracted during the great 'flu epidemic of 1919. The value of antibiotics to health is thus obvious. Add to this the large proportion of agricultural production (around a third) lost to the ravages of insects, weeds and disease, and the continuing need for insecticides, herbicides and fungicides becomes self-evident. The benefits of chemistry to our health (drugs), nutrition (agricultural chemicals and food additives) and quality of life (plastics, dyestuffs, paints, synthetic fabrics, perfumery, chemicals and so on) are legion.

John Emsley has attempted to dispel the myths and perceived problems surrounding everyday chemicals while emphasizing their importance. The first half of the book is almost wholly positive, particularly the first two chapters on perfumes and sweeteners. The secrets of Chanel No. 5, Charlie and other well-known brands of perfume are at least partially revealed, and human scents and aphrodisiacs are not forgotten. All of the main synthetic sweeteners are discussed and he convincingly debunks Yudkin's claim that sucrose is "pure, white and deadly". He also dispels fears about saccharin, and makes clear comparisons between the risks associated with everyday chemicals and the much greater risks associated with smoking or driving.

In a chapter on alcohol, he covers the fundamental chemistry of brewing and gives sound advice, based on knowledge of alcohol metabolism, on how to avoid hangovers. More good news derives from evidence that modest drinkers have, on average, less incidence of coronary heart disease, lower cholesterol levels and fewer colds. Many of the myths about cholesterol, fats and fibre are revealed in the following chapter. I was unaware that Denis Burkitt (of Burkitt's lymphoma fame) rediscovered the beneficial properties of fibre in the late 1960s when he noted that Africans rarely suffered from

colon or rectal cancers because of their high intake of fibre (Hippocrates originally recommended wholemeal bread in 400 BC). The chemistry of successful dieting is also discussed. Drugs make a brief (perhaps too brief) appearance in a chapter on painkillers, with the risks again clearly matched against the benefits. For example, there was an estimated 1-in-25,000 chance of dying from the disgraced (but highly effective) antiarthritis drug Opren, yet a 1-in-60 chance of dying from surgery associated with joint replacement and a 1-in-5 chance of dying from the long-term effects of smoking.

The benefits of chemistry become a little harder to sustain in the succeeding chapters, although Emsley sings the praises of plastics, especially PVC (polyvinyl chloride). He emphasizes the dangers of relying too much on animal tests, exemplified by the enormous disparity in dioxin toxicity between guinea pigs and other mammals. He gives a frank and fair account of all of the accidents involving the release of dioxins, but also reveals that these chemicals are produced in vast amounts in such natural processes as forest fires. But he does not mention the possible long-term effects on fertility of PCBs (polychlorinated biphenyls), DDT (dichlorodiphenyltrichloroethane) and so on, which are only now being revealed.

There is much irrational fear about nitrates in drinking water, and run-off from rivers rarely provides as much dietary nitrate as do green vegetables (up to 3,000 parts per million in spinach) or Perrier water (17 milligrams per litre). But there is no room for complacency: levels of nitrate in rainwater have increased dramatically since the invention of the motor car, which has also increased the concentration of carbon dioxide (and other gases) in the atmosphere. Emsley is convinced that global warming is not due to carbon dioxide, although he does not mention the effects of chlorofluorocarbons on the ozone layer, a much more serious result of our use of chemicals.

Emsley does not shrink from introducing chemical terms and principles where necessary, but the general reader should have no problem with these chemical asides. For those who like to see chemical structures, there is a glossary at the end of the book. □

This accessible and entertaining book deserves to be widely read by the general public, and contains much information that will be of use to those who teach elementary chemistry courses. It may also persuade those journalists and environmentalists who persistently misrepresent chemicals to give chemistry the credit it so richly deserves. □

John Mann is in the Department of Chemistry, University of Reading, PO Box 224, Whiteknights, Reading RG6 2AD, UK.

The evolutionary dynamics of repetitive DNA in eukaryotes

Brian Charlesworth, Paul Sniegowski & Wolfgang Stephan

Repetitive DNA sequences form a large portion of the genomes of eukaryotes. The 'selfish DNA' hypothesis proposes that they are maintained by their ability to replicate within the genome. The behaviour of repetitive sequences can result in mutations that cause genetic diseases, and confer significant fitness losses on the organism. Features of the organization of repetitive sequences in eukaryotic genomes, and their distribution in natural populations, reflect the evolutionary forces acting on selfish DNA.

REPETITIVE DNA sequences form a substantial fraction of the genomes of many eukaryotes¹⁻³. This class includes satellite DNA (very highly repetitive, tandemly repeated sequences), minisatellite and microsatellite sequences (moderately repetitive, tandemly repeated sequences), and transposable elements (moderately repetitive, mobile, dispersed sequences). In many cases these sequences seem to be maintained solely by their ability to replicate within the genome (the 'selfish DNA' hypothesis)^{4,5}. Far from conferring benefits, their behaviour can sometimes result in a fitness loss to the host⁶. Some human genetic diseases are known to be caused in this way, including mutations due to insertions of transposable elements^{7,8}, to chromosomal rearrangements induced by recombination between repeated sequences^{9,10}, or to the amplification of microsatellite sequences¹¹.

It has often been proposed that repetitive sequences are functionally important for the host organism¹²⁻¹⁶, or are maintained because their mutagenic activities contribute to the long-term evolutionary potential of the population^{17,18}. But these may be consequences rather than causes of the presence of repeated sequences. Here we present the main findings of theoretical and empirical work on the evolutionary and population biology of repetitive sequences, with emphasis on how the selfish DNA hypothesis can be tested by comparing the predictions of population genetic models with the data. In addition, we will show that some general features of the organization of eukaryote genomes may be simple consequences of the evolutionary forces acting on selfish DNA.

Tandemly repeated non-coding DNA

There is an extraordinary degree of variation in genome size between different eukaryotes, which bears little relation to differences in organismal complexity, ploidy level, or numbers of genes that code for proteins¹⁻⁵. For example, the newt *Triturus cristatus* has around six times as much DNA as humans, who have about 7.5 times as much as the pufferfish *Fugu rubripes*^{2,19}. Much of this variation is due to non-coding, tandemly repeated sequences, whose basic biology and evolutionary dynamics will be considered next.

Properties of tandem arrays. Variation in the numbers and lengths of tandemly repeated units occurs on many different scales (Box 1). Several genetic mechanisms can affect the numbers of repeating units in tandem arrays, including gene amplification by processes such as replication slippage^{20,21} rolling circle amplification²², unequal exchange^{23,24}, and mutation by base substitutions²⁵ (Fig. 1 and Box 2). The relative importance of these processes, and the values of the parameters that describe them, are still uncertain.

Functional explanations for tandem arrays have been favoured by several workers¹²⁻¹⁴. The *Responder* satellite

sequences of the *Drosophila melanogaster* segregation distorter system clearly affect the fitness of their carriers²⁶. There is also evidence for biological effects of heterochromatin (which is rich in satellite sequences), such as the binding by heterochromatin of certain proteins^{3,13}, and its role in achiasmate chromosome segregation in *Drosophila*²⁷. But these properties of heterochromatin may not have a role in the origin and maintenance of repeated sequences: they could simply be a response to their presence in the genome. Comparative patterns, such as the inverse correlation between development rate and amount of highly repeated DNA in groups such as flowering plants and salamanders^{2,3,28,29}, are compatible with non-selective mechanisms for maintaining repeated arrays. Selection against large amounts of repetitive DNA may vary in intensity with life-history; for example low rates of cell division caused by larger genome size are probably less harmful to organisms that develop relatively slowly. More slowly developing species could therefore accumulate larger quantities of repetitive sequences generated by array size expansion (Box 2). The comparative data seem to fit this permissive interpretation better than the adaptive alternatives which have been proposed²⁹. The large differences in amounts of highly repetitive sequences between closely related species, and the absence of measurable fitness effects of large deletions or duplications of heterochromatin in *Drosophila*, also suggest strongly that selection does not actively maintain the high copy numbers of many of these sequences^{1,3-5}.

The forces affecting simple DNA sequences (microsatellites) are best understood. *In vitro* studies suggest that strand slippage during DNA replication is the major cause of the observed length polymorphism of microsatellites within populations²¹. This process results in length changes of a few bases at a time. There is apparently a relation between replication slippage of microsatellite sequences and defective DNA repair, as mutations in DNA repair genes can cause tract instability when errors resulting from slippage replication remain unrepaired^{30,31}. The sudden increases in the lengths of trinucleotide repeat arrays (up to 10-fold) that cause several human genetic diseases (for example fragile X syndrome) seem, however, to involve other, unknown, mechanisms^{11,32,33}.

Several processes may lead to the expansion or contraction of arrays of minisatellite repeats. At some well-studied human minisatellite loci, a bias towards gains of a few repeats has been found, with gains occurring preferentially at one end of a tandem array³⁴. New repeats on a given chromosome can originate either from the same chromosome or from its homologue. In addition, unequal exchange could occur between tandem arrays on sister chromatids or between arrays on homologous chromosomes^{23,24} (Fig. 1). At present, however, there is no direct evidence for unequal exchanges at minisatellite sequences³⁴. The frequencies with which changes in array size occur at micro- and minisatellite

BOX 1 Major classes of tandemly repeated noncoding sequences

1. **MICROSATELLITE** (simple) sequences: arrays of short (2–5 bp) nucleotide repeats, found in vertebrate, insect and plant genomes⁹⁸. At least 30,000 microsatellite loci are present in the human genome, located in the euchromatin⁹⁹. Copy numbers are characteristically variable within a population, typically with mean array sizes of the order of 100 but with multiple array size classes distributed around the mean^{35,98}.
2. **Minisatellite** sequences: arrays of longer (~15 bp) repeats, generally involving mean array lengths of 0.5–30 kb. They are found in euchromatic regions of the genome of vertebrates, fungi and plants, and are also highly variable in array size⁵⁰.
3. **Satellite** sequences: repeat units can be similar in length (5–10 bp) to micro- or minisatellites, or much larger (~100 bp). They are typically organized as large (up to 100 megabases (Mb)) clusters in the heterochromatic regions of chromosomes, near centromeres and telomeres or on the Y chromosome^{2,3,52,53}. They are apparently not as variable in array size within populations as micro- and minisatellites^{45,46}.

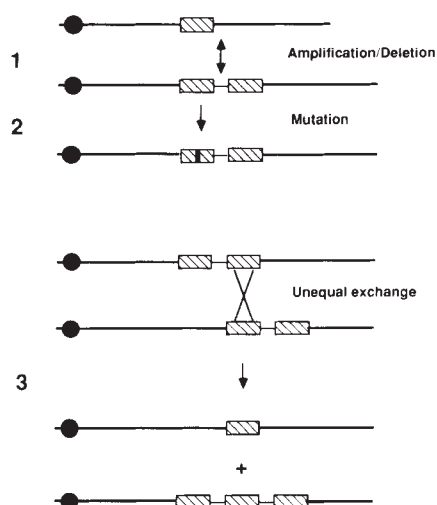


FIG. 1 Processes that may affect tandem repetitive sequences. 1, The effects of amplification and deletion events on the size of an array; 2, the effect of mutation on the composition of an array; 3, the effect of unequal exchange on variation in array size.

loci are much higher than normal mutation rates, sometimes being as high as a few per cent per generation^{34,35}.

The forces governing satellite DNA evolution are less well understood, mainly because the large sizes of satellite DNA clusters preclude direct experimental analysis, and also hinder the collection of meaningful population data. Because replication slippage involves only a short stretch of DNA at a time, it probably plays only a minor part in the amplification of very long satellite repeat units. Slippage may, however, be important in the initial formation of satellite repeats, which are generally composed of smaller subunits^{20,25,36}. The observation of extra-chromosomal circular satellite DNA²² suggests the possibility that long satellite repeats may primarily be amplified by extra-chromosomal rolling circle replication, followed by reinsertion into the genome. The occurrence of unequal exchange has been inferred from sequence analysis of satellites containing higher-order repeat units^{37,38}. The question of whether unequal exchange occurs predominantly between sister chromatids or between homologues is still open, although data on the tandem arrays of the *PSR* supernumerary chromosome of *Nasonia vitripennis* (which is confined to the haploid males) can only be explained by mitotic intrachromosomal or sister-chromatid exchanges³⁹.

BOX 2 Population dynamics of tandem arrays

THE state of a haploid genome with respect to a given class of tandemly repeated DNA sequence at a particular chromosomal location is characterized by the number of repeats in the array, i . The state of the population is described by the distribution of values of i among haploid sets, such that x_i is the frequency of chromosomes with array size i . The mean array size is $\bar{i}$. The following factors affect the change of this distribution over time.

1. **Amplification** (Fig. 1): a sudden increase in array length by one or more repeat units. It may involve replication slippage^{20,21}, rolling circle replication of an extrachromosomal copy of an array followed by reinsertion²², conversion-like processes³⁴, or other, unknown, mechanisms¹¹.

2. **Deletion**: a sudden decrease in array length by one or more units, including replication slippage and unequal intrachromatid exchanges which result in the formation of an extra-chromosomal circle^{22,36}.

Both of these processes can be included in a model in which the probability that an array of length j generates an array of length i is αP_{ij} , where α is the probability of a change in array length per generation, and P_{ij} is the conditional probability of the change from j to i (refs 36, 40, 41, 43, 48, 49).

3. **Unequal exchange** (Fig. 1): this can occur either between homologous chromosomes, between sister chromatids or within chromatids. We will consider only the first category here^{23,48,101}. Let γ be the probability of an exchange and Q_{ijk} be the conditional probability that an exchange produces a daughter chromosome with array size i from parental chromosomes with copy numbers j and k ($0 < i < j+k$). There is an equal chance of production of a daughter with $j+k-i$ copies. The net probability of the production of a daughter chromosome with i or $j+k-i$ copies is then γQ_{ijk} .

4. **Selection**: this is usually assumed to be very weak or non-existent, until array sizes reach a relatively high value, such that fitness declines rapidly to zero with increasing numbers of repeats for individuals carrying more than Ω copies^{48,49}.

5. **Genetic drift**: the frequencies of haploid genomes with different array sizes are subject to multinomial sampling due to finite effective population size N_e , in the same way as a conventional multiple allele system with allele frequencies x_i (refs 40, 41, 43, 48, 49).

6. **Mutation**: mutational differences in sequence among members of an array may affect their probability of undergoing unequal exchange^{24,25,44}.

In an infinite population, a unique non-trivial equilibrium distribution of array size is generated by unequal exchange in the absence of selection²³. In a finite population, this distribution is perturbed by genetic drift^{36,48}. In the absence of amplification, a population that is fixed for chromosomes with array size 1 is trapped there, because each chromosome contains a single copy, so that unequal exchange can no longer generate variability. The expected time $\bar{t}(\bar{i}_0)$, to reach this state, starting from an initial mean copy number $\bar{i}_0$, can be obtained in the following asymptotic parameter ranges⁴⁵:

$$\bar{t}(\bar{i}_0) \sim \begin{cases} \gamma^{-1} \ln(\bar{i}_0/2) & 4N_e\gamma \ll 1 \\ N_e \ln(\bar{i}_0/2) & 4N_e\gamma \gg 1 \end{cases} \quad (1)$$

These results indicate that, for a given population size, the time to absorption at copy number 1 is inversely proportional to γ as long as recombination is so low that $4N_e\gamma \ll 1$. As recombination rate increases, the dependence on γ weakens and then disappears totally. If repeated sequences are occasionally regenerated by amplification processes, the steady-state is such that mean array size is highest in regions of low recombination¹⁰¹.

Selection against high copy numbers in combination with amplification, unequal exchange and genetic drift results in a situation such that mean array size is high only when $N_e\gamma \ll 1$, and is highest when selection against long arrays is weakest¹⁰¹, as might be the case for organisms with long development times^{2,28,29}.

Population processes and data. Population genetic models of tandem arrays generally assume that they are non-functional, that is to say, neutral or slightly deleterious. Base substitutions in the sequence are considered to be selectively neutral, but there may be natural selection against increased array size (Box 2).

The frequency distributions of length alleles at more than 100 microsatellite loci in human populations generally fit a stepwise mutation model quite well. In such a model, arrays increase or decrease in repeat number by one or two repeats, with occasional increases by larger steps^{35,40,41}. The fit of the population data to this model suggests that mutational steps usually involve only very few repeat units, as with replication slippage (Box 2).

In contrast to microsatellites, the length distributions of minisatellite alleles do not fit a neutral stepwise mutation model⁴². This suggests that events that alter copy number occasionally involve many repeat units, as can happen with unequal exchange. On the other hand, the population data are not consistent with a model in which every new array that appears in the population is unique, as would be expected if events such as unequal exchanges with unconstrained misalignments were frequent⁴³. Perhaps the best evidence for the occurrence of unequal exchange at minisatellite loci comes from the observation that minisatellites showing the greatest level of allelic variability in repeat copy number generally have low levels of inter-repeat sequence variability⁴⁴. This homogeneity in tandem arrays is difficult to explain by events which are restricted to certain parts of arrays, such as the gene-conversion-like amplification processes mentioned above³⁴. Instead, forces such as unequal exchange that act over the entire array must be invoked⁴⁴. The observed bias towards gain of repeats³⁴ would cause the expansion of minisatellite arrays over evolutionary time, unless opposed by selection acting against very large arrays, deletions caused by some unknown mechanism, or the interaction between genetic drift and unequal exchange (Box 2).

In contrast to micro- and minisatellites, satellite sequences usually show little variation in copy number within populations, relative to their large array sizes^{45,46} (the satellite sequence of the *Responder* locus in *D. melanogaster*³⁷ seems to be atypical in this respect). However, substantial differences in array size are often observed between isolated populations and closely related species^{3,47}, so that processes causing expansion or contraction of satellite arrays must occur at a significant rate on an evolutionary timescale.

Evolutionary dynamics. Modelling of the evolutionary dynamics of tandem arrays suggests that the salient characteristics of tandem-repetitive non-coding DNA discussed above are inter-related. Unequal exchange in combination with random genetic drift and selection can have a strong effect on the accumulation of tandem-repetitive sequences in the genome (Box 2). In particular, accumulation of very highly repeated sequences is expected to occur only in regions with very low recombination and weak selective constraints on array length^{48,49}. There are reasons why it may be selectively advantageous for recombination to be suppressed near centromeres and telomeres, as is indeed the case in many eukaryotes⁴⁸, and such regions are typically havens for tandem repeats³. It therefore seems likely that suppressed recombination causes the accumulation of repetitive sequences; in addition, sequences which are inherently unlikely to recombine will tend to be more abundant⁴⁸. In contrast, minisatellites, which are less repetitive but much more variable in copy number than satellites, occur in the euchromatic portions of chromosomes, where crossing over is more frequent⁵⁰. In the absence of any deterministic forces that cause an increase in array size, repeats will eventually be lost by genetic drift, with probability one^{36,48,49,51}. As expected if this is the case, there are examples of the turnover of arrays between closely related species⁵¹.

Another characteristic difference between satellite and minisatellite DNAs is the greater tendency for multimeric repeats to be found in satellites. The formation of higher-order repeats is particularly often seen in larger arrays (>200 kilobases (kb))⁵². The best studied example is the alpha satellite DNA family in primates. This consists of a basic repeat unit of 171 base pairs (bp) that is combined into various higher-order structures, ranging from dimers to 16-mers³⁸. Computer simulations incorporat-

ing unequal exchange and selection suggest that higher-order repeats are preferentially formed in regions of low recombination and weak selective constraints on total array size^{25,44}, provided that crossing over between repeats depends on short stretches of sequence identity rather than long segments of overall homology²⁴. Evidence for this has recently been found for satellite DNAs^{37,38}. The simulations also indicate that unequal exchange is a strong long-range ordering force which can keep tandem arrays homogeneous, even if levels of crossing over are low relative to mutation. This may explain why satellites, such as the chromosome-specific human alpha satellite families³⁸, are so homogeneous, even though unequal exchange and gene amplification are relatively rare events. Tandem arrays of similar, relatively short, repeat units (≈ 10 bp) are often juxtaposed in heterochromatic regions⁵³. This pattern is seen in simulations with low rates of unequal exchange and weak selection²⁵.

Transposable elements

Much of the moderately dispersed repeated DNA of eukaryotes appears to consist of transposable elements (TEs), which are sequences capable of inserting copies of themselves into new genomic locations⁵⁴. Transposable element insertions cause a substantial proportion of known mutations with major phenotypic effects⁵⁴, and may also contribute to quantitative trait variation⁵⁵. Their abundances vary considerably across species; for example, *D. melanogaster* appears to have about four times as much middle repetitive DNA as its close relative *D. simulans*⁵⁶.

Properties of transposable elements. Eukaryotic transposable elements are divided into two main classes, according to their mechanism of transposition (see Box 3). Within a given host species, there are usually several different families of elements in each class. Members of the same family generally share considerable sequence homology. Within a large group of species, such as the genus *Drosophila*, it is usually impossible to recognize the presence of a particular family outside a certain taxonomic range^{57,58}. This suggests that families of transposable elements arise and are lost from lineages over evolutionary time. The processes involved are not well understood, although some plausible models have been proposed⁵⁹. There are several examples of a lack of congruence between the presence or absence of an element family and the phylogeny of its host species. This indicates that across-species transfer of transposable elements can occur, although this is probably a rare evolutionary event^{58,60,61}.

The mammalian retroviruses are well known to be infectious (horizontally transmissible) agents, although they are also capable of transposition^{54,58}. The natural horizontal transfer of retrotransposons and other types of transposable element has not been observed, although the *gypsy* element of *D. melanogaster* has been experimentally transferred between individuals⁶². The spread and maintenance of most transposable elements in host populations is probably mediated by vertical (germline) transmission from parent to offspring during sexual reproduction (Fig. 2).

Consequences of transposable element activity. The majority of mutations caused by element insertions and element mediated chromosomal rearrangements must be deleterious, as is the case for other forms of mutation^{6,63}. However, it has been argued that transposable element activity may cause mutations that could have important developmental and evolutionary effects^{15,17,18,64-66}, and that the beneficial consequences of this activity ensures their persistence.

Evidence for this is largely indirect. In some species, the transcription and transposition of some transposable elements is correlated with specific developmental stages⁵⁴, perhaps suggesting functional importance for the host organism¹⁵. Such phenomena, however, could merely represent the parasitic subversion of host functions by transposable elements⁶⁷. The apparently preferential transposition of certain *Drosophila* retrotransposons to chromosome ends is a process which balances

BOX 3 The main classes of eukaryote transposable elements

1. ELEMENTS which transpose through reinstitution of the product of reverse transcription of an RNA copy of the element (retroelements)*

a, Retrotransposons without long terminal repeats, for example *Drosophila* I and R2 elements, mammalian LINEs, trypanosome *ingi* elements^{54,58}. Complete copies of these elements encode proteins needed for reverse transcription, but defective elements are often abundant^{54,75}.

b, Retrotransposons with long terminal repeats, for example *Drosophila copia*, *gypsy*, *Arabidopsis Ta1*, yeast Ty elements^{54,58}.

c, Retroviruses (these have infectious capability, in addition to a requirement for insertion into the host genome as part of their life cycle): for example avian leukosis virus, mammalian leukaemia virus, mammalian immunodeficiency viruses^{54,58}.

d, Other elements that encode proteins needed for reverse transcription: mammalian and avian hepadnaviruses, plant caulimoviruses, group II mobile introns of fungi and plant mitochondria and plastids^{54,58}.

e, Elements that do not encode proteins needed for reverse transcription: for example mammalian SINES⁷⁶.

2. Elements that transpose directly through DNA copies

Elements in this category are characterized by inverted terminal repeats, and open reading frames coding for 'transposase': for example *Drosophila* P and *hobo*, animal *mariner*, maize *Ac*, *Ds* and *Mu*, nematode *Tc1*, sea-urchin *TU* elements⁵⁴.

Details of the mechanism of transposition are not known in most cases. P elements transpose by excision and reinsertion of an element from a single chromatid, followed by filling of the resultant double-strand gap, by copying the element on the sister chromatid¹⁰⁰. The maize *Ac* and *Antirrhinum Tam3* elements transpose by excision during chromosome replication, which may result in a gain of copy number if the element reinserts into a non-replicated chromatid⁵⁴. Defective elements that lack transposase activity, but that can transpose in the presence of transposase encoded by complete elements, are frequently present⁵⁴.

* The above classification is not intended to reflect phylogenetic relationships among retroelements. The origin and evolutionary relationships of the retroelements have been reviewed by Eickbush⁵⁸.

replication-associated telomere loss^{68,69}, but the host genome may simply be exploiting a property of these elements for its own benefit. Experiments with asexual yeast strains indicate that the Ty element insertions can occasionally be beneficial in evolving laboratory populations, even though average fitness declines with increasing Ty copy number⁶⁵. In a sexual population, individual element insertions may rise to high frequency as a result of beneficial insertional mutations, but this will not increase the frequencies of other transposable elements⁷⁰. This process thus cannot explain the spread and maintenance of large numbers of elements in natural populations of sexual organisms, in the face of selection against the majority of insertions.

Chromosomal distribution of transposable elements in natural populations. Surveys of the distribution of transposable elements in genomes from natural populations of *D. melanogaster* have been conducted by several investigators. One approach uses the technique of *in situ* hybridization of transposable element probes to the polytene salivary gland chromosomes. This allows mapping of element insertions to the level of visible polytene chromosome bands (representing 20–200 kb DNA). In a sample of chromosomes from a population, an insertion is typically found only once or a few times at a given location^{59,71,72}; that is, element frequencies are nearly always very low at chromosomal sites into which insertions can occur. This rules out the notion that element insertions frequently induce beneficial mutations, as such insertions would quickly be fixed by selection. Finer-scaled data from restriction mapping studies of defined genomic regions confirm this result, and show that transposable elements almost never occupy coding regions in wild

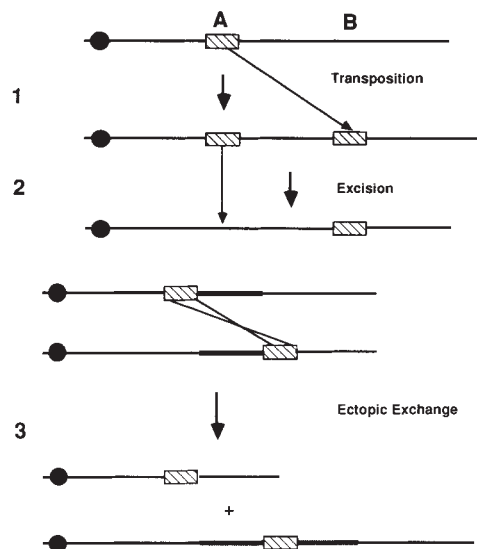


FIG. 2 Processes that may affect the numbers of copies of members of a family of transposable elements. 1 and 2, The effects of transposition and excision, respectively. 3, The effect of ectopic exchange between two elements inserted at different locations on homologous chromosomes; the segment of host DNA that separates the insertion sites (thicker line) is deleted in one exchange product and duplicated in the other.

Drosophila genomes⁵⁹. Because transposable elements are capable of insertion into *Drosophila* coding regions⁵⁴, selection must usually rapidly eliminate such insertions from natural populations.

A limited amount of data is available from *in situ* hybridization surveys in other *Drosophila* species (*D. affinis*, *D. algonquin*, *D. heteroneura* and *D. sylvestris*). These studies suggest higher element abundances and frequencies at occupied salivary chromosome bands for at least some element families^{73,74}. Data on the nucleotide site locations of these insertions are unavailable, however, and thus it is not certain that they are truly at high frequencies. In contrast, restriction mapping studies of other *Drosophila* species, such as *D. pseudoobscura* and *D. ananassae*, suggest element frequencies that are similar to, or even lower than, those found for *D. melanogaster*⁵⁹. Although comparable population surveys have not been conducted in mammals, the highly abundant LINE and SINE elements are evidently often fixed or at high frequencies at individual sites^{75,76}, in contrast to the *Drosophila* findings, although there are some instances of polymorphism for the presence or absence of elements at individual sites^{77–79}.

Transposable element copy number dynamics. The forces that can act to increase or decrease element copy number have been incorporated into population genetic models (Box 4). The *Drosophila* population data indicate that the magnitude of the deterministic forces controlling element abundance is much greater than that of genetic drift (Box 4). The predominantly low populational frequencies of transposable elements imply the existence of deterministic forces that oppose the spread of elements by transposition⁵⁹. The mean copy numbers of most elements seem to be roughly at equilibrium between an increase in copy number under replicative transposition and a decrease in copy number brought about by the other forces listed in Box 4, in accordance with the selfish DNA hypothesis⁵⁹.

An important role for self regulation of transposition in containing the spread of most *Drosophila* transposable elements is ruled out by the observed excess of transposition over excision events in laboratory stocks^{59,80}; this would not be expected if the rate of transposition balances the rate of excision, as is required for equilibrium under self regulation (Box 4). Some

BOX 4 Population dynamics of transposable elements

THE state of a population of diploid individuals with respect to a family of transposable elements is characterized by the distribution of the numbers of genomic copies of these elements among host individuals. With fairly free recombination between chromosomal sites occupied by elements, this distribution will be approximately binomial⁵⁹. The dynamics of the mean number of elements per individual, $\bar{n}$, is affected by the following factors:

1. Transposition: let the probability that a given element present in the germ line of a host individual produces a new copy be u per generation for the germ line. Self regulation of transposition means that u is a decreasing function of the number of elements in the family in question carried by the genome of the host individual, u_n . Such self regulation has been observed for a number of transposable element families that replicate directly through DNA intermediates, such as *P* elements, and also for the retrotransposon *I* in *Drosophila*^{54,59}. It may be an evolved response to deleterious effects of transposition¹⁰² or of element insertions⁵⁹.

2. Excision: let the corresponding probability per generation of loss of an element from its chromosomal site be v . Excision has been observed for elements that replicate directly through DNA intermediates, but is very rare or non-existent for retrotransposons^{54,59,80}.

3. Selection: the simplest model for the effects of natural selection on transposable element abundance assumes that the fitness of an individual with n copies of a given element family is a function of n , w_n . There are at least two potential sources of selection against elements, causing w_n to be a decreasing function of n (ref. 59).

a. Selection against deleterious mutations induced by insertions of transposable elements.

b. Selection against deleterious chromosomal rearrangements created by recombination between elements of the same family that are present in different chromosomal sites (ectopic exchange, see Fig. 2).

With infinite population size and no linkage disequilibrium between sites occupied by elements, and assuming that new element copies insert randomly into unoccupied sites, the change in $\bar{n}$ per generation is given by the approximate expression

$$\Delta \bar{n} \approx \bar{n} \left(1 - \frac{\bar{n}}{2m} \right) \left(\frac{\bar{c} \ln \bar{w}}{\bar{c} \bar{n}} + u_n \right) - \bar{n}v \quad (2)$$

where m is the number of sites available for occupation by members of the element family in a haploid genome, and $\bar{w}$ is the mean fitness of the population (that is, the mean of w_n over all individuals)⁵⁹.

These equations can readily be modified to model the dynamics of defective elements that rely on transposase provided by complete elements^{59,102}.

4. Genetic drift (finite population size): random sampling of element frequencies at individual chromosomal sites can be modelled by standard population genetic methods. These enable the probability distribution of the frequency of an element in the population at a given chromosomal site to be calculated, as a function of the effective size of the local breeding population, N_e , and the deterministic forces 1–3 (refs 59, 95). Fits of observed patterns of element frequencies per site to this distribution suggest that the magnitude of the deterministic forces decreasing copy number ($v - \bar{c} \ln \bar{w} / \bar{c} \bar{n}$) is much greater than $1/N_e$, so that the population behaves deterministically to a good approximation^{59,71}.

Equilibrium conditions are given by putting $\Delta \bar{n} = 0$ in equation (2). In the absence of selection, equilibrium occurs either when all sites are filled up with elements ($\bar{n} = 1/2m$), or when self regulation of the rate of transposition is such that $u_n = v$. The mean frequency per site of elements is equal to $\bar{n}/2m$. A low value of this frequency, as observed in *Drosophila*, requires either that the latter condition is satisfied (self regulation), or that selection is such that w_n is a steeper-than-linear function of n (ref. 59).

When element frequencies are low at each site, the mean fitness of the population at equilibrium, relative to the fitness of an element-free individual, is equal to $\exp(-\bar{n}(u-v))$ (ref. 59). In *Drosophila*, $u-v$ is typically of the order of 10^{-4} or 10^{-5} for most element families, although some elements are capable of moving at much higher rates, and the rates may be influenced by both environmental and genetic factors^{59,80}. The copy number of all element types in a typical wild genome is probably somewhere between 500 and 1,000 (ref. 59). These values imply that $\bar{w}$ is around 0.90–0.95, so that the genetic load associated with *Drosophila* transposable element families is quite low⁵⁹.

For mammalian SINE elements, the assumption that copy numbers and element frequencies are in equilibrium is inappropriate, and stochastic models have to be used⁷⁹.

form of natural selection against elements as a function of their abundance in the genome is therefore required (Box 4). The net force tending to increase the mean abundances of transposable elements is likely to be very weak for most element families, and the total genetic load associated with selection against them in *Drosophila* is accordingly small (Box 4).

Some progress has been made in understanding the possible nature of selection against transposable elements in the *Drosophila* genome. Data on the chromosomal distributions of elements do not suggest an important role for selection against insertional mutations in non-coding regions. In particular, the partial recessivity of most deleterious mutations of minor effect implies more intense selection against X-chromosomal insertions compared with autosomal insertions, as X-chromosomal mutations are fully expressed in males⁸¹. But the expected magnitude of deficiency of elements on the X chromosomes relative to the autosomes is not observed for most element families^{72,81,82}.

Another mechanism for the containment of transposable elements is their role in causing chromosomal rearrangements through recombinational processes^{59,83}. Meiotic exchange between elements at non-homologous sites (ectopic exchange) can cause deletions and duplications of chromosomal material (Fig. 2), which are expected to have dominant deleterious effects in many cases. This can oppose transposable element increase by reducing the fertility of individuals with high element copy numbers (Box 4). There is direct evidence for the occurrence of ectopic exchange between element sequences in *Drosophila*⁸⁴, and in other species including humans^{9,54,85,86}.

The survey data from *D. melanogaster* show that transposable elements are significantly overabundant in chromosomal regions in which the rate of meiotic recombinational exchange is reduced, as expected if ectopic exchange is reduced in parallel with regular meiotic exchange^{59,82,83}. In addition, transposable elements are often overabundant within and around recombination-suppressing chromosomal inversions, relative to uninverted homologous regions, in both laboratory and natural populations^{81,87,88}. Many families of *Drosophila* elements are also abundant in the β -heterochromatin, a region of restricted recombination^{3,89}. The one inconsistency in this pattern is the absence of any detectable transposable element increase in the distal region of the X chromosome, where normal meiotic recombination is highly reduced^{59,82}.

Associations of transposable elements with genomic areas of reduced recombination have been observed in groups other than *D. melanogaster*. In *D. miranda*, the non-recombining neo-Y chromosome has accumulated a large excess of transposable element sequences^{90,91}. Similarly, the Y chromosomes of *D. hydei* and of mice carry many copies of retrotransposons^{92,93}. *In situ* hybridization studies of the genomes of four species of deer mice (*Peromyscus*) have revealed disproportionately high concentrations of retrotransposons on the X and Y chromosomes, for which meiotic exchange is reduced relative to the autosomes⁹⁴. While these findings are consistent with the ectopic exchange model, the absence of selection against insertional mutations in genetically inert regions such as the Y chromosome, and the stochastic accumulation of deleterious elements in regions with no recombination^{95,96}, could also have a role.

Overview

Population genetic studies of transposable elements in *Drosophila* provide strong support for models of such elements as selfish DNA. Few population studies have been done in other taxa, except in *Saccharomyces cerevisiae* in which *Ty* element location seems to be quite variable among strains^{65,97}, consistent with the *Drosophila* findings. The contrast between the patterns for *Drosophila* and (probably) yeast elements, with their predominantly low frequencies at individual sites, and those for the mammalian SINE and LINE elements, which seem to show fixation or near-fixation at most chromosomal sites, is a major question for theoretical and empirical research. At present, we can only speculate about the reasons for this difference⁸³. Further information on mechanisms by which the spread of transposable elements is opposed is needed, particularly for species other than *D. melanogaster*.

The nature of the forces controlling the array sizes of tandem repeats also requires further research. Progress will probably come both from increased knowledge of the molecular basis of array size expansion and contraction, and from the comparison of patterns of within-population variation in array size with models based on this knowledge. A promising start on the latter has been made. The discovery of the causative role of the expansion

of trinucleotide repeat arrays in several important human genetic diseases, and the fact that the starting point for this amplification may be arrays of large size within the normal range of variation^{11,33}, has revealed an unexpected practical significance to what might seem a purely academic problem.

Finally, the population genetic models of both transposable elements and tandem arrays show that the dynamics of these sequences are strongly affected by the frequency of genetic recombination. For rather different reasons, both types of repeats are likely to accumulate differentially in genomic regions where recombination is infrequent. This may explain the long-standing puzzle of the association of repetitive sequences with the non-recombining, heterochromatic regions of many eukaryote genomes^{48,59,96}. Careful empirical scrutiny of the assumptions on which these models are based offers the best hope for assessing their utility. □

B. Charlesworth is at the Department of Ecology and Evolution, University of Chicago, 1101 E. 57th Street, Chicago, Illinois 60637, USA, P. Sniegowski is at the Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48824, USA, and W. Stephan is at the Department of Zoology, University of Maryland, College Park, Maryland 20742, USA.

- Dover, G. A. & Flavell, R. B. (eds) *Genome Evolution* (Academic, London, 1982).
- Cavalier-Smith, T. (ed.) *The Evolution of Genome Size* (John Wiley, Chichester, 1985).
- John, B. & Miklos, G. L. G. *The Eukaryote Genome in Development and Evolution* (Allen and Unwin, London, 1988).
- Doolittle, W. F. & Sapienza, C. *Nature* **284**, 601–603 (1980).
- Orgel, L. E. & Crick, F. H. C. *Nature* **284**, 604–607 (1980).
- Mackay, T. F. C. *Genet. Res.* **48**, 77–87 (1986).
- Wallace, M. R. et al. *Nature* **353**, 864–865 (1991).
- Holmes, S. E., Dombrowski, B. A., Krebs, C. M., Boehm, C. D. & Kazanian, H. G. *Nature Genet.* **7**, 143–148 (1994).
- Lehrman, M. A., Goldstein, D. L., Russell, D. W. & Brown, M. S. *Cell* **48**, 827–835 (1987).
- Lakich, D., Kazanian, H. H., Antonarakis, S. E. & Gitschier, J. *Nature Genet.* **5**, 236–241 (1993).
- Kuhl, D. P. A. & Caskey, C. T. *Curr. Opin. Genet. Dev.* **3**, 404–407 (1993).
- Britten, R. J. & Davidson, E. H. *Science* **165**, 349–357 (1969).
- Brutlag, D. L. A. *Rev. Genet.* **14**, 121–144 (1980).
- Cavalier-Smith, T. in *The Evolution of Genome Size* (ed. Cavalier-Smith, T.) 105–184 (John Wiley, Chichester, 1985).
- McDonald, J. F. *Curr. Opin. Genet. Dev.* **3**, 855–864 (1993).
- Britten, R. J. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5992–5996 (1994).
- Nevers, P. & Saedler, H. *Nature* **268**, 109–115 (1977).
- Syvanen, M. A. *Rev. Genet.* **18**, 271–293 (1984).
- Brenner, S. et al. *Nature* **366**, 265–268 (1993).
- Levinson, G. & Gutman, G. A. *Molec. Biol. Evol.* **4**, 203–221 (1987).
- Schlötterer, C. & Tautz, D. *Nucleic Acids Res.* **20**, 211–215 (1992).
- Okumura, K., Kiyama, R. & Oishi, M. *Nucleic Acids Res.* **15**, 7477–7489 (1987).
- Krüger, J. & Vogel, F. *J. molec. Evol.* **4**, 201–247 (1975).
- Smith, G. P. *Science* **191**, 528–535 (1976).
- Stephan, W. *Molec. Biol. Evol.* **6**, 198–212 (1989).
- Wu, C.-I. & Hammer, M. F. in *Evolution at the Molecular Level* (eds Selander, R. K., Clark, A. G. & Whittam, T. S.) 177–203 (Sinauer, Sunderland, MA, 1991).
- Hawley, R. S. & Theurkauf, W. E. *Trends Genet.* **9**, 310–317 (1993).
- Bennett, M. D. *Proc. R. Soc. Lond. B* **249**, 109–135 (1972).
- Pagel, M. & Johnstone, R. A. *Proc. R. Soc. Lond. B* **249**, 119–124 (1992).
- Strand, M., Prolla, T. A., Liskay, R. M. & Petes, T. D. *Nature* **365**, 274–276 (1993).
- Lindahl, T. *Curr. Biol.* **4**, 249–251 (1994).
- Fu, Y.-H. et al. *Cell* **67**, 1047–1058 (1991).
- Zhong, N., Dobkin, C. & Brown, W. T. *Nature Genet.* **5**, 248–253 (1993).
- Jeffreys, A. J. et al. *Nature Genet.* **6**, 136–145 (1994).
- Di Rienzo, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3166–3170 (1994).
- Walsh, J. B. *Genetics* **115**, 553–567 (1987).
- Cabot, E. L., Doshi, P., Wu, M.-L. & Wu, C.-I. *Genetics* **135**, 477–487 (1993).
- Warburton, P. E., Wayne, J. S. & Willard, H. F. *Molec. cell. Biol.* **13**, 6520–6529 (1993).
- Reed, K. M., Beukeboom, L. W., Eickbush, D. G. & Werren, J. H. *J. molec. Evol.* **38**, 352–362 (1994).
- Edwards, A., Hammond, H. A., Jin, L., Caskey, C. T. & Chakraborty, R. *Genomics* **12**, 241–253 (1992).
- Valdes, A. M., Slatkin, M. & Freimer, N. B. *Genetics* **133**, 737–749 (1993).
- Deka, R., Chakraborty, R. & Ferrell, R. E. *Genomics* **11**, 83–92 (1991).
- Harding, R. M., Boyce, A. J., Martinson, J. J., Flint, J. & Clegg, J. B. *Genetics* **135**, 911–922 (1993).
- Stephan, W. & Cho, S. *Genetics* **136**, 333–341 (1994).
- Jabs, E. W., Goble, C. A. & Cutting, G. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 202–206 (1989).
- Wewrick, R. & Willard, H. F. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9394–9398 (1989).
- Bachmann, L. & Sperlich, D. *Molec. Biol. Evol.* **10**, 647–659 (1993).
- Charlesworth, B., Langley, C. H. & Stephan, W. *Genetics* **112**, 947–962 (1986).
- Stephan, W. *Genet. Res.* **47**, 167–174 (1986).
- Armour, J. A. L. & Jeffreys, A. J. *Curr. Opin. Genet. Dev.* **2**, 850–856 (1992).
- Gray, I. C. & Jeffreys, A. J. *Proc. R. Soc. Lond. B* **243**, 241–253 (1991).
- Tyler-Smith, C. & Willard, H. F. *Curr. Opin. Genet. Dev.* **3**, 390–397 (1993).
- Lohe, A. R., Hilliker, A. J. & Roberts, P. A. *Genetics* **134**, 1149–1174 (1993).
- Berg, D. E. & Howe, M. M. (eds) *Mobile DNA* (American Society for Microbiology, Washington, DC, 1989).
- Mackay, T. F. C., Lyman, R. F. & Jackson, M. S. *Genetics* **130**, 315–332 (1992).
- Dowsett, A. P. & Young, M. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4570–4574 (1982).
- Martin, G., Wiernasz, D. & Schedl, P. *J. molec. Evol.* **19**, 203–213 (1983).
- Eickbush, T. H. in *Evolutionary Biology of Viruses* (ed. Morse, S. S.) 121–157 (Raven Press, New York, 1994).
- Charlesworth, B. & Langley, C. H. A. *Rev. Genet.* **23**, 251–87 (1989).
- Kidwell, M. G. *Genetica* **86**, 275–286 (1992).
- Robertson, H. *Nature* **362**, 241–245 (1993).
- Kim, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1285–1289 (1994).
- Crow, J. F. & Simmons, M. J. in *The Genetics and Biology of Drosophila* (eds Ashburner, M., Carson, H. L. & J. N. Thomson) 1–35 (Academic Press, London, 1983).
- McClintock, B. *Science* **226**, 792–800 (1984).
- Wilke, C. M., Maimier, E. & Adams, J. *Genetica* **86**, 55–73 (1992).
- Shapiro, J. A. *Genetica* **86**, 99–111 (1992).
- Bingham, P. M. & Zachar, Z. in *Mobile DNA* (eds Berg, D. E. & Howe, M. M.) 485–502 (American Society for Microbiology, Washington, D. C. 1989).
- Biessmann, H. et al. *EMBO J.* **11**, 4459–4469 (1992).
- Levis, R. W., Ganesan, R., Houtchens, K., Tolar, L. A. & Sheen, F. *Cell* **75**, 1083–1093 (1993).
- Leigh, E. G. *Genetics* **73** (suppl.), 1–18 (1973).
- Charlesworth, B., Lapid, A. & Canada, D. *Genet. Res.* **60**, 103–114 (1992).
- Blémont, C. *Genetica* **86**, 67–84 (1992).
- Hunt, J. A., Bishop, J. G., Carson, H. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7146–7150 (1984).
- Hey, J. *Molec. Biol. Evol.* **6**, 66–79 (1989).
- Hutchinson, C. L. A., Hardies, S. C., Loer, D. D., Shehee, W. R. & Edgell, M. H. in *Mobile DNA* (eds Berg, D. E. & Howe, M. M.) 593–617 (American Society for Microbiology, Washington, DC, 1989).
- Deininger, P. L. in *Mobile DNA* (eds Berg, D. E. & Howe, M. M.) 619–636 (American Society for Microbiology, Washington, DC, 1989).
- Bellis, M. et al. *Molec. Biol. Evol.* **4**, 351–363 (1987).
- Batzler, M. A. & Deininger, P. L. *Genomics* **9**, 481–487 (1991).
- Tachida, H. & Iizuka, M. *Genetics* **133**, 1023–1030 (1993).
- Nuzhdin, S. V. & Mackay, T. F. C. *Molec. Biol. Evol.* (in the press).
- Montgomery, E. A., Charlesworth, B. & Langley, C. H. *Genet. Res.* **49**, 31–41 (1987).
- Charlesworth, B., Lapid, A. & Canada, D. *Genet. Res.* **60**, 115–130 (1992).
- Langley, C. H., Montgomery, E. A., Hudson, R. R., Kaplan, N. L., Charlesworth, B. *Genet. Res.* **52**, 223–235 (1988).
- Montgomery, E. A., Huang, S.-M., Langley, C. H. & Judd, B. H. *Genetics* **129**, 1085–1098 (1991).
- Mueller, M. W., Allmaier, M., Eskes, R. & Schweyen, R. J. *Nature* **366**, 174–176 (1993).
- Seligm, C. H., Lecellier, G. & Belcour, L. *Nature* **366**, 176–178 (1993).
- Eanes, W. F., Wesley, C. & Charlesworth, B. *Genet. Res.* **59**, 1–9 (1992).
- Sniegowski, P. D. & Charlesworth, B. *Genetics* **137**, 815–827 (1994).
- Vaury, C., Bucheton, A. & Pelisson, A. *Chromosoma* **98**, 215–224 (1989).
- Steinemann, M. & Steinemann, S. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7591–7595 (1992).
- Ganguly, R., Swanson, K. D., Ray, K. & Krishnan, R. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1340–1344 (1992).
- Hochstenbach, R., Harhangi, H., Shouren, K. & Henning, W. *J. molec. Evol.* (in the press).
- Eicher, E. M., Hutchison, K. W., Phillips, S. J., Tucker, P. K. & Lee, B. K. *Genetics* **122**, 181–192 (1989).
- Wichman, H. A., Van Den Bussche, R. A., Hamilton, J. J. & Baker, R. J. *Genetica* **86**, 287–294 (1992).
- Charlesworth, B. in *Population Genetics and Molecular Evolution* (eds Ohta, T. & Aoki, K.) 213–232 (Springer, Berlin, 1985).
- Charlesworth, B. *Science* **251**, 1030–1033 (1991).
- Cameron, J. R., Loh, E. Y. & Davis, R. W. *Cell* **16**, 739–751 (1979).
- Bruford, M. W. & Wayne, R. K. *Curr. Opin. Genet. Dev.* **3**, 939–943 (1993).
- Weissenbach, J. *Curr. Opin. Genet. Dev.* **3**, 414–417 (1993).
- Gloor, N. B., Nassif, N., Johnson-Schlitz, D. M., Preston, C. R. & Engels, W. R. *Science* **1110**–1117 (1991).
- Stephan, W. *Genet. Res.* **50**, 41–52 (1987).
- Brookfield, J. F. Y. *Genetics* **128**, 471–486 (1991).

ACKNOWLEDGEMENTS. We thank D. Charlesworth, C. Langley, J. Werren and Chung-I Wu for their comments on the manuscript. This work was supported by grants from the NIH and NSF.

An early haematopoietic defect in mice lacking the transcription factor GATA-2

Fong-Ying Tsai^{*†}, Gordon Keller[‡], Frank C. Kuo[§], Mitchell Weiss^{*}, Jianzhou Chen^{||}, Margery Rosenblatt^{*}, Frederick W. Alt^{†||} & Stuart H. Orkin^{*†¶}

^{*} Division of Hematology-Oncology, Children's Hospital and Dana Farber Cancer Institute, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

[†] Howard Hughes Medical Institute, Children's Hospital, ^{||} Department of Genetics, Harvard Medical School and Center for Blood Research, [§] Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

[‡] The National Jewish Center, Denver, Colorado 80206, USA

Blood cell development relies on the expansion and maintenance of haematopoietic stem and progenitor cells in the embryo. By gene targeting in mouse embryonic stem cells, we demonstrate that the transcription factor GATA-2 plays a critical role in haematopoiesis, particularly of an adult type. We propose that GATA-2 regulates genes controlling growth factor responsiveness or the proliferative capacity of early haematopoietic cells.

LINEAGE commitment and differentiation of haematopoietic cells occur continuously throughout the life of the individual. At the top of a developmental hierarchy lie rare stem cells capable of both self-renewal and subsequent progenitor expansion. Below, committed precursors undergo terminal differentiation to produce mature blood cells. Within the mammalian embryo the site of haematopoiesis shifts from its initial position in the yolk sac blood islands (primitive haematopoiesis) to the fetal liver and then to the bone marrow (definitive haematopoiesis). The origin of haematopoietic stem cells that populate the adult is controversial. Although it has been proposed that they arise within the yolk sac and migrate to the fetal liver¹, evidence in support of an intraembryonic origin has also been reported^{2,3}.

Haematopoiesis is controlled through the cooperative effects of growth factors that permit cellular proliferation or differentiation, and nuclear transcription factors that activate lineage-specific genes and initiate, or arrest, cell cycling. Three members of the GATA family of transcription factors have emerged as candidate regulators of gene expression in haematopoietic cells⁴. GATA-1, which is expressed at high levels in erythroid cells, mast cells and megakaryocytes^{5,6}, is essential for normal primitive and definitive erythropoiesis^{7,9}. Two related proteins, GATA-2 and GATA-3, are expressed in selected haematopoietic cell populations, as well as in some non-haematopoietic cell types including embryonic stem (ES) cells¹⁰. Whereas GATA-3 is restricted to T-lymphoid cells, GATA-2 is highly expressed in haematopoietic progenitors, early erythroid cells, mast cells and megakaryocytes^{11,14}. In chicken erythroid progenitors, forced expression of GATA-2 (but not GATA-1 or GATA-3) promotes proliferation at the expense of differentiation¹⁵. In *Xenopus* and zebrafish, GATA-2 is expressed in regions of the embryo fated to be involved in haematopoiesis (ref. 16, and H. W. Detrich *et al.*, manuscript in preparation). Taken together, these indirect findings suggest that GATA-2 may serve an important role in haematopoietic development.

This possibility is addressed here by targeted gene disruption in mouse ES cells. We show that GATA-2 is essential for normal mouse development. In its absence mice die by embryonic days 10–11 (E10–E11) with severe anaemia. Our analysis of chimaeras generated with GATA-2 null ES cells and *in vitro* differentiation

of GATA-2 null ES cells revealed a profound deficit in definitive haematopoietic stem or progenitor cells, largely related to poor expansion of the cellular pool in response to haematopoietic growth factors.

Disruption of the mouse GATA-2 gene

We isolated GATA-2 complementary DNA, characterized the mouse genomic locus and then constructed a targeting vector in which a neomycin-resistance cassette replaced the carboxyl zinc-finger domain of the GATA-2 gene (Fig. 1a). Because GATA proteins bind DNA as monomers and the carboxyl zinc finger is essential for DNA binding¹⁷, the targeted replacement constitutes a null mutation. Following selection in G418 and ganciclovir, heterozygous, targeted ES cell clones (GATA-2^{+/-}) of the lines CCE (ref. 18) and J1 (ref. 19) were obtained at frequencies of 10–16% and ~5%, respectively. Representative Southern blot analysis of heterozygous clones is shown in Fig. 1b.

ES cells homozygous for a disrupted GATA-2 gene were isolated by growth in an elevated concentration of G418 (ref. 20). Southern blot analysis revealed only the mutated allele in GATA-2^{-/-} cells (Fig. 1b). To confirm that GATA-2^{-/-} cells do not express an intact GATA-2 transcript, we carried out reverse-transcription polymerase chain reaction (RT-PCR) analysis of differentiating embryoid bodies derived from wild-type and GATA-2^{-/-} cells (Fig. 1c). No intact GATA-2 messenger RNA was detected in GATA-2^{-/-} cells. Thus, the engineered mutation renders these cells unable to encode functional GATA-2 protein even if an aberrant mRNA is produced from the locus.

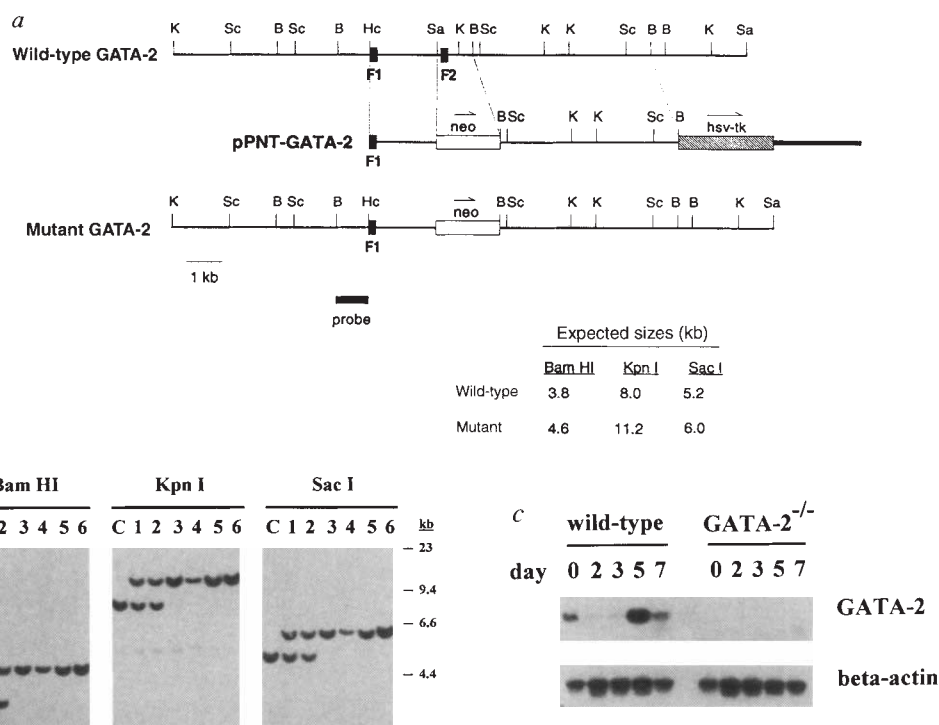
GATA-2^{-/-} embryos die with severe anaemia

Three J1 GATA-2^{+/-} clones achieved germ-line transmission. F₁ mice heterozygous for the GATA-2 gene mutation appeared normal. Heterozygous F₁ animals from two independent ES cell clones were intercrossed to generate F₂ progeny. No mice homozygous for the GATA-2 gene mutation survived to birth among 73 F₂ progeny analysed (Table 1).

To determine the time at which the GATA-2 gene mutation is lethal, E9.5–E11.5 embryos were examined. No GATA-2^{-/-} embryos survived beyond E11.5 (Table 1), and at E10.5, most (67%) GATA-2^{-/-} embryos were dead. These non-viable embryos had enlarged pericardial sacs but otherwise appeared normal. The vessels of the yolk sac, on the other hand, appeared very pale and nearly empty.

¶ To whom correspondence should be addressed.

FIG. 1 Disruption of GATA-2 gene in mouse embryonic stem cells. *a*, Restriction maps of mouse GATA-2 genomic fragment, targeting construct and predicted structure of targeted GATA-2 allele. Restriction enzymes: B, *Bam*HI; K, *Kpn*I; Sc, *Sac*I; Sa, *Sall*; Hc, *Hinc*II. Exons containing the two zinc fingers (F1 and F2) are indicated as solid boxes. The targeting construct pPNT-GATA-2 was derived from pPNT plasmid³⁵. The arrows show the orientation of the neomycin and HSV-tk selection cassettes. Dashed lines denote the borders of homology between wild-type GATA-2 gene and pPNT-GATA-2. A 1.0 kilobase (kb) *Bam*HI/*Hinc*II fragment was used as probe to distinguish the GATA-2 wild-type and mutant alleles in Southern blots. The predicted sizes of restriction fragments from wild-type and mutant alleles are shown. *b*, Southern blot analysis of ES cell clones. Genomic DNA isolated from wild-type CCE cells (C), heterozygous GATA-2 mutants (GATA-2^{+/-} cells, lanes 1 and 2) and homozygous GATA-2 mutants (GATA-2^{-/-} cells, lanes 3–6) were digested with *Bam*HI, *Kpn*I, or *Sac*I and probed as shown in *a*. *c*, Expression of RNA transcripts in differentiating embryoid bodies (EBs) of wild-type and GATA-2^{-/-} ES cells. RNAs were analysed by RT-PCR from day 0 to 7 days of EB development. **METHODS.** Murine GATA-2 genomic clones were isolated from a λ DashII strain 129 library (Stratagene) by standard procedures³⁶. A 1.8-kb *Hinc*II/*Sall* fragment and a 5-kb *Bam*HI fragment were subcloned into pPNT to generate the targeting construct. Hence, a 1-kb region of the gene



encoding F2 was deleted. The construct was electroporated into CCE¹⁸ and J1 (ref. 19) cells, and selected with neomycin and ganciclovir. GATA-2^{-/-} clones were identified by Southern blot analysis. GATA-2^{-/-} ES cells were derived from GATA-2^{+/-} cells by culture in 1.2 mg ml⁻¹ active G418 (ref. 20). For mRNA analysis, 1/10 of the cDNA resulting from reverse transcription of 1 μ g of total RNA was used in PCR reactions using GATA-2 and β -actin primers⁹.

On E9.5, GATA-2^{-/-} embryos were viable. These embryos had an appropriate number of somites (20–25) and were comparable in size and development to littermates. However, the embryo and yolk sac were noticeably pale, allowing unequivocal identification of homozygous mutants (Fig. 2*a, b*). Microscopic examination failed to reveal additional anatomical abnormalities. Mutant embryos had an apparently normal heart, well-formed neural tubes, otic vesicles, two branchial arches, vitelline vasculature and intact fetal membranes (Fig. 2*c, d*). Endothelium, a tissue in which GATA-2 is thought to regulate several genes¹¹, seemed normal (Fig. 2*e, f*). Although the blood of these embryos contained apparently normal primitive erythroid cells (Fig. 2*g, h*), the total number of blood cells from embryo and yolk sac bleeding was reduced two- to sevenfold (average threefold) in homozygotes. Thus, despite apparently normal maturation of red cells, GATA-2^{-/-} embryos show marked anaemia and fail to survive beyond the stage of primitive haematopoiesis.

Haematopoietic defect of GATA-2^{-/-} cells *in vivo*
Contribution of GATA-2^{-/-} cells to haematopoiesis in the early embryo. As GATA-2^{-/-} embryos die before the fetal liver stage, we investigated a requirement for GATA-2 in haematopoiesis by examining chimaeras generated by injecting GATA-2^{-/-} ES cells into wild-type blastocysts. The contribution of GATA-2^{-/-} cells to tissues of chimaeric embryos was assayed at E13 and E17 to evaluate primitive and definitive haematopoiesis, respectively.

Tissue DNA of embryos was assayed for chimaerism by a PCR assay that distinguishes wild-type and mutant GATA-2 alleles. As blood samples of E13 embryos contain predominantly nucleated, primitive erythrocytes of yolk sac origin, the extent of GATA-2 mutant DNA in blood reflects the primitive erythropoietic potential of GATA-2^{-/-} cells. Five embryos, estimated to be 30–70% chimaeric based on PCR with whole-embryo DNA, were analysed at E13. The GATA-2 mutant allele was

TABLE 1 Genotype and yolk sac haematopoietic progenitors of offspring derived from heterozygous parents

Stage	Total (litters)	Genotype			Survival	Colony no. per +/+ or +/- yolk sac			Colony no. per -/- yolk sac		
		+/+	+/-	-/-		Mix	E	M	Mix	E	M
Postnatal	73 (13)	21	52	0	0						
E11.5	39 (4)	9	13	8	0						
E10.5	50 (6)	11	30	9	3	106	444	306	2	6	16
E9.5	58 (6)	19	25	14	14	99	116	222	1	1	37

Male chimaeras were bred with C57BL/6 female mice to generate F₁ GATA-2^{+/-} mice. Genotyping of F₂ offspring from intercrossing GATA-2^{+/-} mice was done on tail DNA at postnatal stage and embryo or yolk sac DNA at embryonic stage by genomic Southern blot or PCR analysis. Surviving embryos were defined as those with beating hearts. Haematopoietic colony assays were performed using cells from collagenase-treated yolk sac as described³⁴. Growth factors included in the methylcellulose were erythropoietin (2 U ml⁻¹), 1.3% pokeweed mitogen-stimulated spleen cell-conditioned medium, and *kit* ligand (100 μ g ml⁻¹). Colonies (>30 cells in size) were scored on the days 7 and 9 of incubation. Under the culture conditions used here, the colonies from E9.5 are of definitive origin. Mix, erythroid–myeloid mixed colonies; E, erythroid colonies; M, myeloid colonies.

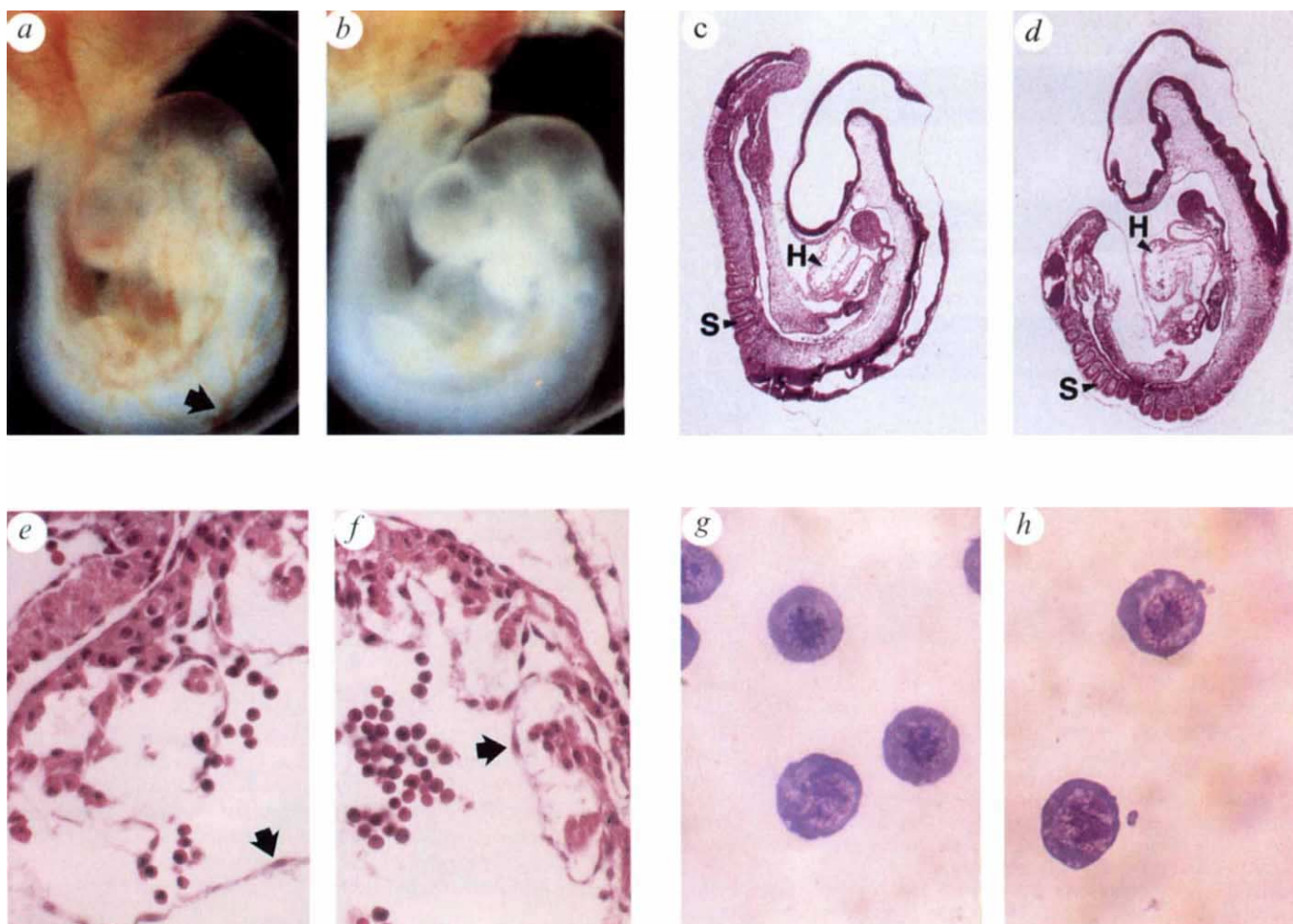


FIG. 2 Phenotypes of wild-type and $GATA-2^{-/-}$ embryos and yolk sac at E9.5. Wild-type (a) and $GATA-2^{-/-}$ (b) E9.5 embryos with intact visceral yolk sac are shown. The arrow points to a branching vitelline vessel on the visceral yolk sac which is red in wild-type and noticeably pale in the mutant. c, and d, Sagittal sections of wild-type (c) and $GATA-2^{-/-}$ (d) E9.5 embryos. H, heart; S, somites. e, f, Sections of developing heart from wild-type (e) and $GATA-2^{-/-}$ (f) E9.5 embryos. Arrows indicate normal-appearing endothelial cells. g, h, May-Grunwald-Giemsa-stained blood cells from wild-type (g) and $GATA-2^{-/-}$ (h) yolk sacs.

METHODS. E9.5 embryos with visceral yolk sac and placenta were

dissected free of maternal tissue and parietal yolk sac with fine forceps. After photography, yolk sac and embryo were detached from placenta and placed in 1 ml of media. Embryos were removed, fixed in 10% neutral formalin, embedded in paraffin, sectioned at 5–7 μ m and stained with haematoxylin and eosin for histological examination. The yolk sac was teased finely to release most of the blood cells into the media. 100 μ l of the cell suspension was spun onto a cytocentrifuge slide and stained with May-Grunwald-Giemsa for morphological assessment; 10 μ l was counted in a haematocytometer to estimate the total blood cell count.

detected in the blood of all embryos, though at a level below that in the whole embryo (Fig. 3). Livers of these embryos were also tested. At this stage, the liver functions primarily as a haematopoietic organ. The contribution of $GATA-2^{-/-}$ cells in fetal livers was virtually undetectable, and far below that in blood samples (Fig. 3).

To assess contribution at the later fetal liver stage, E17 chimeric embryos were examined. At this time circulating red blood cells are non-nucleated; hence, DNA in blood samples reflects the white blood cell fraction. As shown in Fig. 3, $GATA-2^{-/-}$ cells contributed strongly to brain, weakly to fetal liver, but not to blood. The weak contribution to fetal liver at this stage probably reflects emerging hepatic cells rather than haematopoietic cells. This was confirmed by examining *in vitro* haematopoietic progenitors of E13–E18 fetal livers. Among 62 EM colonies (colonies with erythroid and myeloid cells), none was derived from $GATA-2^{-/-}$ cells (data not shown).

After birth, the haematopoietic compartments (blood, bone marrow, thymus and spleen) of chimaeras 1 week to 3 months of age contained no $GATA-2^{-/-}$ cells (Fig. 3). Thus, $GATA-2^{-/-}$ cells are impaired with respect to their ability to contribute to haematopoiesis, particularly during and after the fetal liver stage.

Haematopoietic compartment of the adult mouse. The extent of the deficit in definitive haematopoiesis was examined in additional $GATA-2^{-/-}$ /wild-type chimaeras at 1–3 months of age. The contribution of mutant ES cells to various tissues was determined by genomic DNA Southern blot analysis. As shown for a representative chimaera (Fig. 4a), $GATA-2^{-/-}$ cells, detected by the presence of a fragment diagnostic of the mutant $GATA-2$ allele, failed to contribute to sites of haematopoiesis (bone marrow, spleen, thymus and blood cells), whereas they contributed to all other organs tested. These findings were confirmed in four chimaeras generated with $GATA-2^{-/-}$ clones derived from two independent $GATA-2^{+/-}$ clones. In contrast, $GATA-2^{+/-}$ cells contributed to all tissues, including haematopoietic sites (Fig. 4a). We conclude, therefore, that $GATA-2^{-/-}$ ES cells display a selective deficit in definitive haematopoiesis.

Mature red blood cells. DNA analysis of blood samples detects only the contribution of nucleated cells in the circulation. To examine the contribution of $GATA-2^{-/-}$ cells to mature erythroid cells, we examined the haemoglobin of chimaeras. CCE ES cells of strain 129/SVJ are homozygous for the Hbb^d β -globin haplotype, whereas host blastocysts of strain C57BL/6 are homozygous for Hbb^s (ref. 7). Hence, the relative ratio of Hbb^d to Hbb^s in blood samples reflects chimaerism in circulating

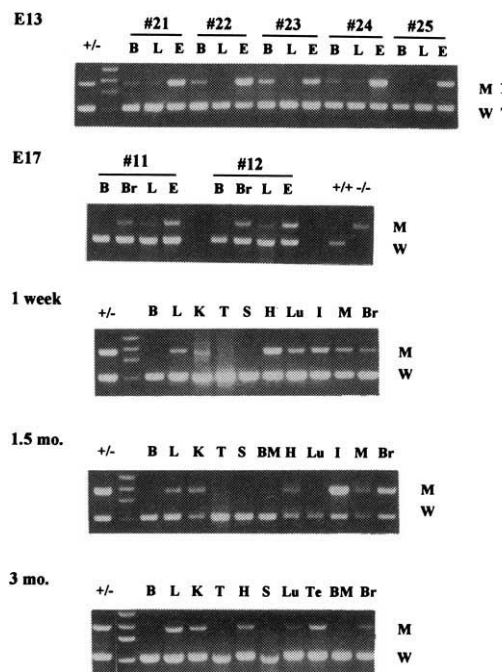


FIG. 3 Tissue contribution of GATA-2^{-/-} cells in GATA-2^{-/-}/wild-type chimaeras at various developmental stages. PCR was used to detect the GATA-2 wild-type (W) and mutated (M) alleles in tissues of chimaeras killed at E13 (embryos 21–25), E17 (embryos 11 and 12), 1 week, 1.5 months and 3 months. Blood (B), brain (Br), liver (L), whole embryo (E), kidney (K), thymus (T), spleen (S), bone marrow (BM), heart (H), lung (Lu), intestine (I), muscle (M) and testis (Te) were examined. Control DNA samples (+/+), (+/-) and (-/-) were prepared from wild-type, GATA-2^{+/+} and GATA-2^{-/-} ES cells, respectively. METHODS. Two independent CCE-derived GATA-2^{-/-} clones were introduced into MF1 or C57BL/6 blastocysts to generate chimaeras as described³⁷. Genomic DNA was isolated from various organs of chimeric embryos or mice and analysed by PCR using allele-specific primers. The targeted band (M, ~950 base pairs, bp) is generated with primers: neo503, GATCTCCTGTCTCATCTCACCTTGCT; and mGATA208, GCTGGACATCTTCGATTCCGGGT; the wild-type band (W, ~600 bp) is generated with primers: mGATA39, GGAACGCCAACGGGGAC; and mGATA208.

red blood cells. Seven out of eight chimaeras generated with GATA-2^{-/-} ES cells (estimated level of chimaerism 30–80% based on coat colour) showed significant levels of *Hbb*^d, compared with none of six chimaeras generated with GATA-2^{-/-} ES cells (estimated level of chimaerism 50–80%) (Fig. 4b). Thus, GATA-2^{-/-} ES cells also fail to contribute to mature, circulating red blood cells *in vivo*.

Myelo-erythroid progenitors in bone marrow and spleen. Although the total bone marrow and spleen of GATA-2^{-/-}/wild-type chimaeras failed to show a significant contribution by the GATA-2^{-/-} cells, we investigated whether haematopoietic colonies derived from the mutant cells could be detected in haematopoietic progenitor assays *in vitro*. Bone marrow and spleen cells of GATA-2^{+/+}/wild-type and GATA-2^{-/-}/wild-type chimaeras were collected and assayed in semi-solid medium for erythroid, myeloid and mixed colonies. The origin of individual haematopoietic colonies was determined by a PCR assay. Of 204 colonies from bone marrow and 87 colonies from spleen cells obtained from 8 GATA-2^{-/-}/wild-type chimaeras, none arose from the GATA-2^{-/-} ES cells (data not shown). In contrast, individual colonies of GATA-2^{+/+}/wild-type chimaeras were either of wild-type or GATA-2^{+/+} origin, as anticipated (data not shown). Thus, GATA-2^{-/-} ES cells do not contribute to the bone marrow and spleen progenitor population of adult chimaeras. This finding is consistent with haematopoietic colony assays of E9.5 yolk sac cells of GATA-2^{-/-} embryos, in which the number of EM-mixed colonies was 1.5% of normal (Table 1).

Lymphoid system. The lymphopoietic potential of GATA-2^{-/-} ES cells was examined by RAG-2-deficient (recombination activating gene-2) blastocyst complementation²¹ and found to be markedly reduced. As RAG-2^{-/-} mice produce no mature B or T lymphocytes, mature lymphocytes in chimaeras generated by injection of GATA-2^{-/-} ES cells into RAG-2^{-/-} blastocysts must be derived from the donor ES cells. Of three GATA-2^{-/-}/RAG-2^{-/-} chimaeras analysed, all had low levels of mature T cells in spleen, but only chimaera 3 had a significant number of T cells in the thymus (Fig. 4c). With respect to B-cell lymphopoiesis, all three chimaeras had low levels of serum IgM (data not shown), but no detectable mature B cells in spleen and bone marrow (Fig. 4c). Southern blot analysis (Fig. 4d) revealed a contribution of GATA-2^{-/-} cells to the spleen only in chimaera 3 in which some T cells were present. There was no detectable contribution to bone marrow in any of the three chimaeras. In non-haematopoietic compartments, GATA-2^{-/-} cells contributed to all seven tissues tested, with lower levels in lung and intestine, as shown by Southern blot (Fig. 4d) and PCR analyses (data not shown). Strong selection for mature lymphoid cells in the RAG-2-minus environment accounts for the presence of low numbers of lymphoid cells in thymus or spleen of chimaeras in the face of an undetectable contribution of GATA-2^{-/-} ES cells to bone marrow. Reduced lymphopoiesis from GATA-2^{-/-} cells reflects a lesion at the early progenitor or stem cell level, as committed lymphoid cells do not express detectable GATA-2 (data not shown).

Taken together, our results reveal a cell-autonomous defect in the haematopoietic potential of GATA-2^{-/-} cells. The haematopoietic consequences of GATA-2 loss are most extreme for definitive haematopoiesis, though a deficit in primitive erythropoiesis is evident *in vivo*. All lineages are ultimately affected in adult chimaeras, most probably due to a failure of GATA-2^{-/-} stem cells or progenitors to compete with host stem cells or progenitors during expansion of the haematopoietic system.

In vitro differentiation of GATA-2^{-/-} ES cells

In vitro haematopoietic development^{22,23} of wild-type and mutant ES cells was used to characterize the nature of the deficits of GATA-2^{-/-} cells further. Embryoid bodies (EBs) generated from ES cells were disaggregated at selected days and the cells replated into semi-solid medium containing various combinations of growth factors^{22,23}. Analysis at different times permits quantification of precursors for specific haematopoietic lineages as they develop within the EBs in a defined kinetic pattern. The first detectable haematopoietic precursors to arise are those of the primitive erythroid lineage (Ery^P), a population comparable to the cells present in the mouse yolk sac. As shown in Fig. 5, Ery^P precursors develop within GATA-2^{-/-} EBs, although their frequency is markedly less than that found in GATA-2^{+/+} (data not shown) or wild-type EBs. Shortly after the appearance of the Ery^P precursors, precursors for the definitive erythroid (Ery^D) and macrophage (Mac) lineages develop within EBs. Ery^D precursors are scored following growth in c-kit ligand (KL) plus erythropoietin (Epo) or in a combination of growth factors (KL/Epo/interleukins (IL) 1 and 3/granulocyte-macrophage colony-stimulating factor (GM-CSF)). KL/Epo-responsive GATA-2^{-/-} definitive precursors were absent in day 6 or 10 EBs (Fig. 5). Replating of day 10 EB cells with the combination of factors revealed some GATA-2^{-/-} Ery^D colonies, albeit at a reduced frequency. Macrophage precursors also develop within GATA-2^{-/-} EBs, at a reduced frequency (Fig. 5). Mast cell precursors, which appear later in EB development, respond strongly to KL, but not significantly to IL-3/IL-1 (data not shown). As with the Ery^D precursors, there is a striking absence of KL-responsive mast cell precursors in GATA-2^{-/-} EBs, yet some colony formation in the presence of KL/IL-3/IL-1 (Fig. 5). A similar pattern of precursor development was observed in two independent GATA-2^{-/-} clones; GATA-2^{+/+} cells displayed wild-type behaviour (not shown). In summary, *in vitro*

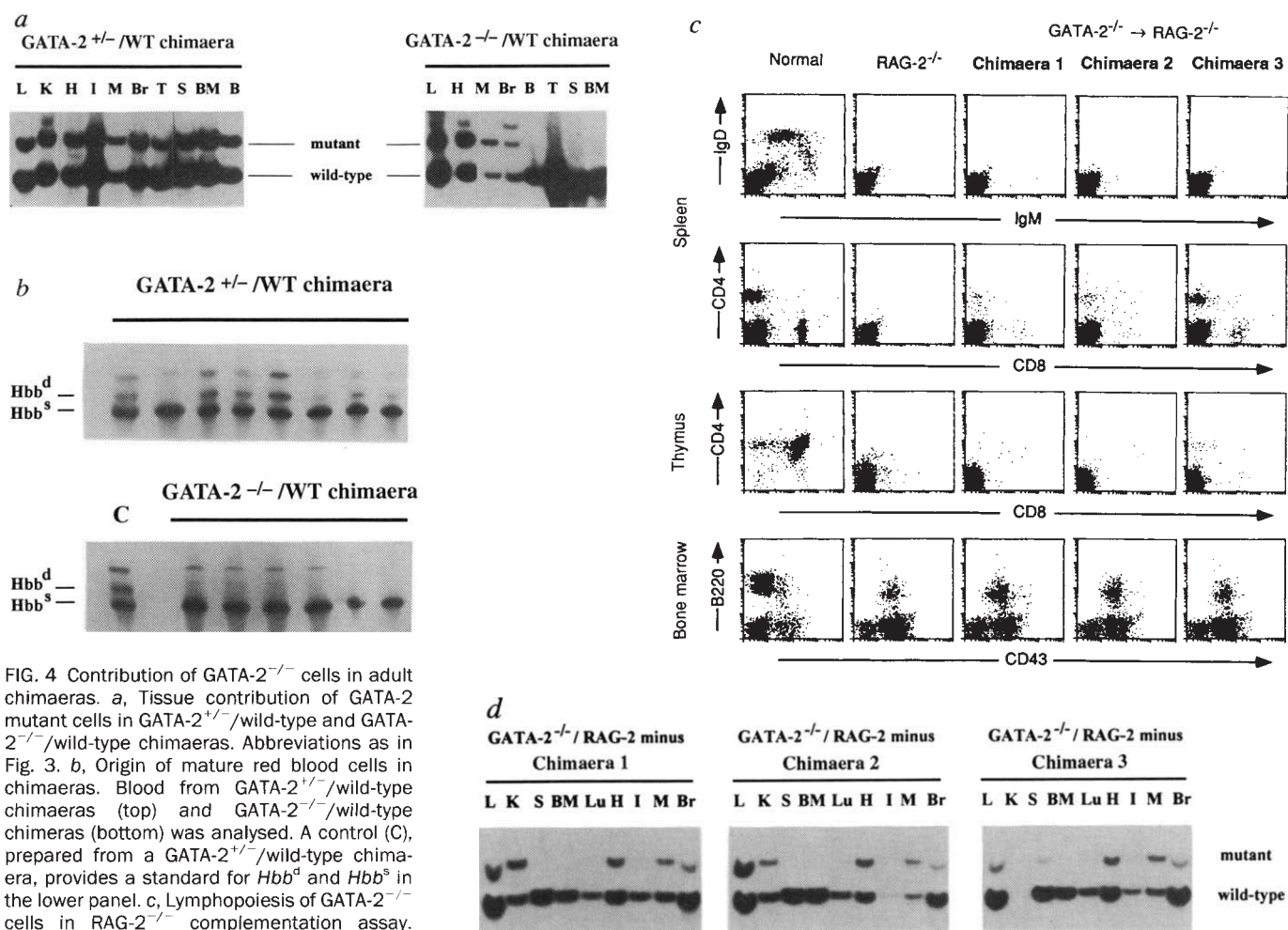


FIG. 4 Contribution of GATA-2^{-/-} cells in adult chimaeras. **a**, Tissue contribution of GATA-2 mutant cells in GATA-2^{+/-}/wild-type and GATA-2^{-/-}/wild-type chimaeras. Abbreviations as in Fig. 3. **b**, Origin of mature red blood cells in chimaeras. Blood from GATA-2^{+/-}/wild-type chimaeras (top) and GATA-2^{-/-}/wild-type chimaeras (bottom) was analysed. A control (C), prepared from a GATA-2^{+/-}/wild-type chimaera, provides a standard for Hbb^d and Hbb^s in the lower panel. **c**, Lymphopoiesis of GATA-2^{-/-} cells in RAG-2^{-/-} complementation assay. Dual-channel fluorescence-activated cell sorting analysis on mononuclear cells from spleen, thymus and bone marrow of three chimaeras, a normal control and a RAG-2^{-/-} animal are shown. **d**, Tissue contribution of GATA-2^{-/-} cells in GATA-2^{-/-}/RAG-2^{-/-} chimaeras. Genomic DNA isolated from various tissues of GATA-2^{-/-}/RAG-2^{-/-} chimaeras was analysed by Southern blot analysis. **METHODS.** Genomic DNA isolated from tissues of chimaeras was analysed by Southern blot analysis as in Fig. 1. Haemoglobin analysis on whole blood from chimaeras was as described³⁸. For the RAG-2

complementation experiment, GATA-2^{-/-}/RAG-2^{-/-} chimaeras were generated and analysed as described^{21,30}. Mononuclear cells from spleen, thymus and bone marrow were examined by flow cytometry analysis using fluorescein (fl)- and phycoerythrin (PE)-linked antibodies. Spleen cells were stained with fl-anti-IgD and PE-anti-IgM; or fl-anti-CD8 and PE-anti-CD4. Thymocytes were stained with fl-anti-CD8 and PE-anti-CD4. Bone marrow cells were stained with fl-anti-B220 and PE-anti-CD43.

differentiation analysis of GATA-2^{-/-} ES cells reveals a profound deficiency of KL-dependent definitive erythroid or mast cell colonies and reduced formation of primitive erythroid and macrophage colonies. As colony formation improved somewhat for both KL-dependent cell lineages in the presence of enhanced growth factor combinations, these data point to a selective failure of GATA-2^{-/-} haematopoietic progenitors to respond to KL.

Discussion

Our results reveal a pivotal role for GATA-2 in haematopoiesis. Primitive haematopoiesis in the yolk sac of mutant mice is sufficiently reduced to lead to severe anaemia and death by E10–E11. Definitive haematopoiesis, that which is evident in fetal liver and thereafter maintained in spleen and bone marrow of adults, is profoundly affected (>50-fold) from GATA-2^{-/-} cells. Although we cannot distinguish between impairment at the stem or progenitor cell level, the loss of virtually all lineages *in vivo* argues for an effect very early in the haematopoietic hierarchy. Our data are most consistent with GATA-2 serving as a regulator of genes controlling growth-factor responsiveness or proliferative capacity of early haematopoietic cells.

The proposed involvement of GATA-2 in early aspects of haematopoiesis is consistent with its expression during embryogenesis (ref. 16, and H.W. Detrich *et al.*, manuscript in preparation) and in haematopoietic cell maturation^{13,14}. The presence of a functional GATA site in the GATA-1 promoter²⁴ raises the formal possibility that GATA-2 might activate expression of the GATA-1 gene in a simple hierarchical pathway. This is not the case, as GATA-2^{-/-} erythroid cells collected from *in vitro* differentiation of ES cells express approximately normal levels of GATA-1 mRNA (our unpublished data). On the contrary, it seems that GATA-1 represses GATA-2 expression during erythroid maturation⁹. As GATA-1 and GATA-2 share some aspects of their function and overlap in their expression, it is plausible that the residual primitive haematopoiesis evident from GATA-2^{-/-} cells *in vivo* and *in vitro* reflects partial compensation because of the presence of GATA-1.

The phenotype of GATA-2^{-/-} embryos is distinct from other mouse mutations affecting haematopoiesis. Severe varieties of dominant white spotting (*W*) and steel (*Sl*) mutants²⁵, gene targeted *c-myb*^{-/-} mice²⁶ and mice deficient in the retinoblastoma gene (*Rb*^{-/-})^{27–29} die at mid-fetal liver stage (~E15–16) with anaemia. *c-myb*^{-/-} and *Rb*^{-/-} mice show no haemato-

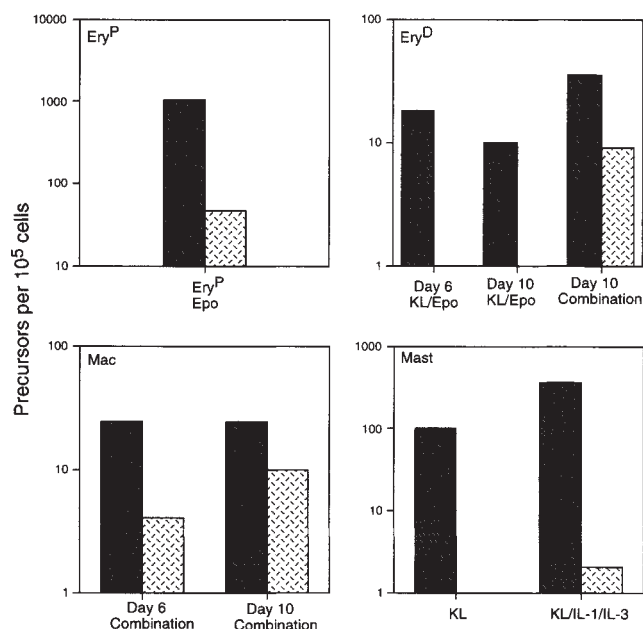


FIG. 5 *In vitro* differentiation of wild-type and GATA-2^{-/-} ES cells. Primitive erythroid (Ery^P), definitive erythroid (Ery^D), macrophage (Mac) and mast cell colonies (Mast) were scored in a two-step replating assay as described²³. Wild-type and GATA-2^{-/-} ES cells were allowed to form embryoid bodies (EBs) in methylcellulose medium. EBs were then dissociated at 6 (Ery^P, Ery^D, Mac), 10 (Ery^D, Mac) or 14 (Mast) days and replated into secondary cultures containing the indicated growth factors. Combination includes IL-1 (1,000 U ml⁻¹), IL-3 (1% conditioned medium by cells transfected with an IL-3-expressing vector³⁹, Epo (2 U ml⁻¹), KL (100 ng ml⁻¹ recombinant or 1% conditioned medium by KL-transfected CHO cell line), GM-CSF (20 U ml⁻¹), G-CSF (1,000 U ml⁻¹), M-CSF (100 U ml⁻¹), IL-6 (5 ng ml⁻¹), IL-11 (25 ng ml⁻¹). Wild-type represented by dark bars, GATA-2^{-/-} by light bars.

poietic defect at the yolk sac stage. A proliferative defect of definitive haematopoietic progenitors seems to be the primary lesion in *c-myb*^{-/-}. The erythropoietic defect seen in *Rb*^{-/-} mice seems to be due in part to abnormalities in the haematopoietic microenvironment³⁰. The closest parallels of GATA-2 deficiency are *W* and *Sl* mice, which result from mutations in the *c-kit* receptor and *c-kit* ligand (or stem cell factor), respectively³¹. Homozygotes for severe *W* or *Sl* alleles die of anaemia by E15 (refs 25, 32). An effect of *W* mutations on yolk sac erythropoiesis has been inferred from a reduction in circulating red blood cells at E13 (ref. 33). The haematopoietic impairment in GATA-2^{-/-} cells or mice is different in that it becomes evident at an early embryonic stage with marked anaemia by E9.5 and affects all lineages.

As GATA-2^{-/-} cells cannot efficiently populate the haematopoietic system, we infer that GATA-2 normally controls genes that encode proteins pivotal for proliferation and/or maintenance of early progenitors or stem cells. From *in vitro* differentiation experiments we conclude that there is no strict block to either erythroid or mast cell development in GATA-2^{-/-} cells, as colony formation improves when precursors are given a broad array of growth factors. The GATA-2^{-/-} precursors studied *in vitro* display a weak response to *kit* ligand compared with other growth-factor combinations. However, expression of *c-kit* mRNA in GATA-2^{-/-} cells (unpublished data) excludes the simple possibility that loss of GATA-2 secondarily leads to *c-kit* deficiency. The defect in GATA-2^{-/-} cells seems to extend beyond the *c-kit*/KL pathway, and may involve multiple signalling systems.

Grossly normal non-haematopoietic development to E9.5 in GATA-2^{-/-} embryos implies that other early developmental processes are independent of GATA-2 or proceed normally

through the action of a related transcription factor, such as GATA-3, which is often coexpressed. Our experiments do not reveal the roles of GATA-2 in other tissues beyond E9.5. GATA-2 is expressed at high levels in the central nervous system and peripheral sympathetic neurons after E11 (J. D. Engel, personal communication) and also in endothelial cells of adult origin¹¹. To access these later developmental stages it may be necessary to rescue embryos by haematopoietic transplantation, carry out chimera experiments using host blastocysts from mice with other targeted mutations in genes affecting these other lineages, or perform tissue-specific gene targeting.

In summary, the experiments described here establish a second member of the GATA factor family, GATA-2, as an essential nuclear regulator for haematopoietic cells. Our findings have immediate implications for the biology of haematopoietic stem cells. First, use of GATA-2 gene expression as a marker may assist in establishing the origin of haematopoietic stem cells in the developing embryo. Second, identification of authentic GATA-2 target genes in early definitive haematopoietic cells should define components of growth-factor response systems critical to their proliferation and/or survival. Third, study of growth factors and cellular interactions that induce or regulate GATA-2 gene expression in early haematopoietic cells may provide a link between environmental influences and expansion of haematopoietic progenitors *in vivo*. Finally, as a pivotal regulator of haematopoietic cell proliferation, the GATA-2 gene represents a candidate target gene for mediating the leukaemogenic effects of transcription factors aberrantly expressed as a result of chromosomal translocations. □

Received 9 May; accepted 20 July 1994.

- Moore, M. S. A. & Metcalf, D. *Br. J. Haemat.* **18**, 279–296 (1970).
- Medvinsky, A. L., Samoylina, N. L., Muller, A. M. & Dzierzak, E. A. *Nature* **364**, 64–66 (1993).
- Godin, I. E., Garcia-Porrero, J. A., Coutinho, A., Dieterlen-Lievre, F. & Marcos, M. A. R. *Nature* **364**, 67–70 (1993).
- Orkin, S. H. *Blood* **80**, 575–581 (1992).
- Tsai, S. F. et al. *Nature* **339**, 446–451 (1989).
- Evans, T. & Felsenfeld, G. *Cell* **58**, 877–885 (1989).
- Pevny, L. et al. *Nature* **349**, 257–260 (1991).
- Simon, M. C. et al. *Nature Genet.* **1**, 92–98 (1992).
- Weiss, M. J., Keller, G. & Orkin, S. H. *Genes Dev.* **8**, 1184–1197 (1994).
- Yamamoto, M. et al. *Genes Dev.* **4**, 1650–1662 (1990).
- Dorfman, D. M., Wilson, D. B., Bruns, G. A. & Orkin, S. H. *J. Biol. Chem.* **267**, 1279–1285 (1992).
- Visvader, J. & Adams, J. M. *Blood* **82**, 1493–1501 (1993).
- Leonard, M., Brice, M., Engel, J. D. & Papayannopoulou, T. *Blood* **82**, 1071–1079 (1993).
- Mouthon, M.-A. et al. *Blood* **81**, 647–655 (1993).
- Briegleb, K. et al. *Genes Dev.* **7**, 1097–1109 (1993).
- Kelley, C., Yee, K., DeCaprio, J., Harland, R. M. & Zon, L. I. *Dev. Biol.* (in the press).
- Martin, D. & Orkin, S. *Genes Dev.* **4**, 1886–1898 (1990).
- Robertson, E. J., Bradley, A., Kuehn, M. & Evans, M. *Nature* **323**, 445–448 (1986).
- Li, E., Bestor, T. H. & Jaenisch, R. *Cell* **69**, 915–926 (1992).
- Mortensen, R. M., Conner, D. A., Chao, S., Geisterfer-Lowrance, A. A. T. & Seidman, J. G. *Molec. Cell. Biol.* **12**, 2391–2395 (1992).
- Chen, J., Lansford, R., Stewart, V., Young, F. & Alt, F. W. *Proc. natn. Acad. Sci. U.S.A.* **90**, 4528–4532 (1993).
- Wiles, M. V. & Keller, G. *Development* **111**, 259–267 (1991).
- Keller, G., Kennedy, M., Papayannopoulou, T. & Wiles, M. V. *Molec. Cell. Biol.* **13**, 472–486 (1993).
- Tsai, S.-F., Strauss, E. & Orkin, S. H. *Genes Dev.* **5**, 919–931 (1991).
- Russell, E. S. *Adv. Genet.* **20**, 358–420 (1979).
- Mucenski, M. L. et al. *Cell* **65**, 677–689 (1991).
- Lee, E. Y.-H. L. et al. *Nature* **359**, 288–294 (1992).
- Jacks, T. et al. *Nature* **359**, 295–300 (1992).
- Clarke, A. R. et al. *Nature* **359**, 328–330 (1992).
- Chen, J. et al. *Curr. Biol.* **3**, 405–413 (1993).
- Fleischman, R. A. *Trends Genet.* **9**, 285–290 (1993).
- Geissler, E. N., McFarland, E. C. & Russell, E. S. *Genetics* **97**, 337–361 (1981).
- Russell, E. S., Thompson, M. W. & McFarland, E. C. *Genetics* **58**, 259–270 (1968).
- Wong, P. M. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3851–3854 (1986).
- Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning, a Laboratory Manual* 2nd edn, (Cold Spring Harbor Laboratory Press, New York, 1989).
- Robertson, E. *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (IRL Press, Oxford, 1987).
- Whitney, J. B. *Biochem. Genet.* **16**, 667–672 (1978).
- Karasuyama, H. & Melchers, F. *Eur. J. Immun.* **18**, 97–104 (1988).

ACKNOWLEDGEMENTS. G.K. and F.C.K. contributed equally to this work. We thank S. Goff for assistance in characterization of GATA-2 genomic clones; V. Stewart and L. Davidson for generation of RAG-2 chimaeras; Genetics Institute for providing a KL-producing cell line, IL-11, and M-CSF; Immunex for providing recombinant KL; and Hoffmann-LaRoche for recombinant IL-1. J.C. is supported by the Cancer Research Institute fellowship and the Arthritis Investigator Award. F.W.A. and S.H.O. are Investigators of the HHMI.

A carbon–oxygen star as progenitor of the type Ic supernova 1994I

K. Nomoto*, H. Yamaoka†, O. R. Pols‡, E. P. J. van den Heuvel§, K. Iwamoto*, S. Kumagai|| & T. Shigeyama*

* Department of Astronomy, School of Science, University of Tokyo, Tokyo 113, Japan

† Department of Physics, Faculty of Science, Kyushu University, Fukuoka 810, Japan

‡ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

§ Astronomy Institute and Centre for High Energy Astrophysics, University of Amsterdam, Amsterdam, The Netherlands

|| Institute of Physical and Chemical Research, Wako, Saitama 350-01, Japan

SUPERNOVAE are classified observationally as type I (which have no hydrogen emission lines in their optical spectra) or type II (which do have hydrogen lines), and are further subdivided¹ into types Ia, Ib, Ic, IIP, IIL and IIb. Type II supernovae are generally thought to result from the explosion of massive stars, and type Ia supernovae from white-dwarf stars that have accreted sufficient mass from a companion to trigger explosion^{1,2}. The origins of types Ib and Ic are much less clear. The Ic class has been particularly controversial, with two main alternatives: the explosion of a massive star that has lost its hydrogen and helium envelopes through fast stellar winds, exposing the carbon- and oxygen-rich core; or the transfer of material to a companion, leaving a small, bare carbon–oxygen (C+O) star which subsequently explodes^{3,4}. We show here that the observed characteristics of SN1994I (in the nearby galaxy NGC5194), which has recently^{5,6} been classified as type Ic, are best understood on the basis of the second of the two alternatives above: that the progenitor was in a close binary system and lost its outer envelopes, leaving a C+O star (of mass ≈ 2 solar masses) which exploded when its iron core collapsed.

To account for the lack (or weakness) of hydrogen and helium features in type Ic supernovae, massive (C+O)-rich stars have been suggested for the progenitors⁴. The difficulty with such massive-star models may be that the decline of the calculated light curve is significantly slower than the observations of type Ic supernovae, even for the reduced ejecta mass down to 2.7 solar masses ($2.7M_{\odot}$; ref. 7). The relatively fast decline of the type Ic light curves can be reproduced by the explosions of low-mass helium stars if extensive mixing of ^{56}Ni to the surface layer occurs^{8,9}. For these low-mass helium-star models, however, theoretical spectra predict (contrary to observations) strong He lines due to radioactive excitation from $^{56}\text{Ni} \rightarrow ^{56}\text{Co}$ decays, thus suggesting that type Ic supernovae are deficient in helium^{10,11}. These difficulties have led to the suggestion that C+O stars that have lost even their He envelope are the progenitors of type Ic supernovae^{11–13}.

Here we propose three evolutionary paths for the formation of C+O star progenitors, and then compare a theoretical light curve of a C+O star explosion with the observations of SN 1994I and other supernovae.

In a massive binary, a bare C+O star can form after two stages of mass transfer¹⁴. The first mass transfer occurs when the more massive star has formed a helium core, and its envelope expands to the Roche lobe (the volume inside which material is gravitationally bound to the star). The H-rich envelope is lost and a helium star is produced. After core He burning, the helium star expands and may again fill its Roche lobe, depending on its remaining mass M_a . This second mass transfer is more likely to occur for lower-mass helium stars because they attain larger radii¹⁵; for example, the maximum radii of helium stars with

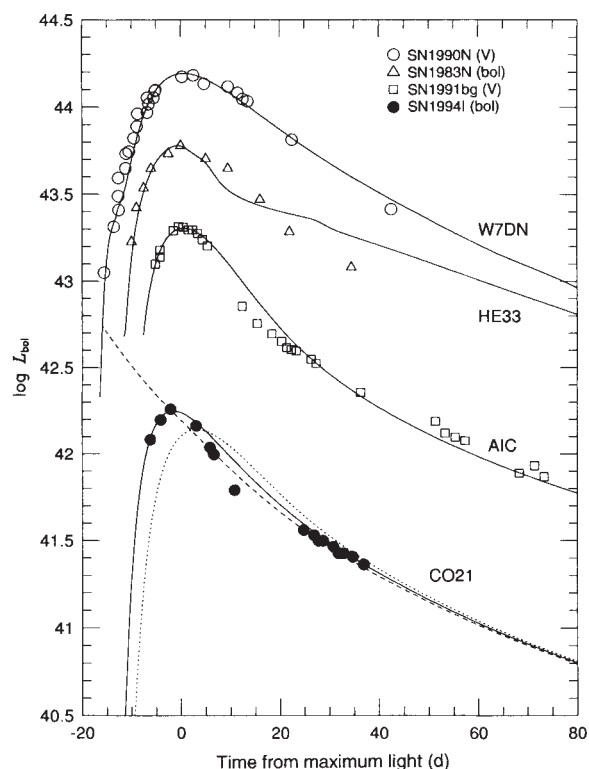


FIG. 1 The bolometric light curves of model CO21 (ejecta mass $M_{\text{ej}} = 0.88M_{\odot}$) are shown for optical opacity $\kappa = 0.1$ (solid line) and 0.2 (dotted line). The dashed line shows the γ -ray energy deposition rate. Other solid curves are bolometric light curves of the $3.3M_{\odot}$ helium-star model (HE33: $M_{\text{ej}} = 2.1M_{\odot}$, $M(^{56}\text{Ni}) = 0.15M_{\odot}$, no mixing⁸), the thermonuclear explosion of a white dwarf (W7DN: $M_{\text{ej}} = 1.4M_{\odot}$, $M(^{56}\text{Ni}) = 0.63M_{\odot}$, $\kappa = 0.3$)²⁶, and the white-dwarf collapse model (AIC: $M_{\text{ej}} = 0.7M_{\odot}$, $M(^{56}\text{Ni}) = 0.15M_{\odot}$, $\kappa = 0.3$)²⁷, where $\log L_{\text{bol}}$ is shifted by 1.2, 1.2 and 0.8, respectively. Filled and open symbols show observed light curves of type Ic SN1994I (bolometric; B. Schmidt, personal communication), type Ib SN1983N (bolometric²⁸), type Ia SN1990N (visual²⁹) and a peculiar type Ia SN1991bg (visual³⁰), where $\log L$ is appropriately shifted except for SN1994I.

$M_a = 3.3, 4.0, 6.0$ and $8.0M_{\odot}$ are 3.7, 3.0, 1.9 and 1.3 solar radii ($R_{\odot}$), respectively¹⁶, and helium stars less massive than $2.7M_{\odot}$ even expand to red-giant dimensions¹⁵.

When the second mass transfer occurs to a more massive companion, it will be conservative (most of the transferred mass is accreted by the companion), and the He star probably retains part of its envelope. On the other hand, if the companion is less massive, the helium star may lose its entire envelope and produce a bare C+O star. In the following, we adopt a value of $2.2M_{\odot}$ for the minimum mass of a He star to undergo core collapse¹⁵.

We propose the following three different evolutionary paths for the formation of C+O stars.

Path A. Initially the binary system consists of star 1 and star 2 whose main-sequence masses are $M_1 \approx (11\text{--}16)M_{\odot}$ and $M_2 \approx (1\text{--}4)M_{\odot}$, respectively, that is, their mass ratio q is between ~ 0.1 and ~ 0.25 . Because of the extreme mass ratio, the first mass transfer is highly non-conservative. This almost inevitably leads to the formation of a common envelope¹⁷ and the subsequent spiral-in¹⁸ of star 2 towards the core of star 1. In many cases, the spiral-in may cause the companions to merge into a single star, but if the initial orbital separation is large enough, the binary system probably survives as a helium star (star 1) of $(2.2\text{--}4)M_{\odot}$ and a main-sequence star (star 2) of $(1\text{--}4)M_{\odot}$, in a much closer orbit. The helium star (1) subsequently expands and undergoes a second, non-conservative mass transfer onto star 2, losing its He envelope in the process. Star 1 then becomes an almost bare C+O star with a low-mass main-sequence companion.

Path B. Star 1 is initially massive enough ($M_1 \geq 11M_\odot$) to evolve through core collapse, leaving a neutron star. The initial mass ratio is not too small, $q \geq 0.4$, so that mass transfer from star 1 to star 2 is (quasi)conservative and star 2 eventually becomes more massive than $11M_\odot$, and subsequently evolves to fill its Roche lobe. Star 1 is now a compact star, and as it is much less massive than star 2, it will inevitably spiral into the envelope of star 2. As in path A, the binary may either merge, or it may survive the spiral-in. The resulting binary system then consists of a helium star (2) of $(2.2\text{--}4)M_\odot$ and compact star (1) of $1.4M_\odot$ in a close orbit. The helium star (2) expands to fill its Roche lobe and undergoes another non-conservative mass transfer to star 1, resulting in a C+O star (2) with a neutron star companion.

Path C. Star 1 has an initial mass in the range $\sim(6\text{--}11)M_\odot$ and evolves to become a C+O or ONeMg white dwarf. The initial mass ratio is close to unity, so that the mass transfer from star 1 to star 2 is almost conservative, and star 2 becomes more massive than $11M_\odot$. Further evolution is similar to path B, except that star 1 is a compact white dwarf. Star 2 undergoes twice non-conservative mass transfer and eventually becomes a bare C+O star with a white-dwarf companion (star 1).

The resultant C+O star (formed through one of these paths) explodes as a type Ic supernova. (If a substantial He envelope is retained in the second mass transfer, the explosion may be observed as a type Ib supernova.)

We can make a rough estimate of the formation rate of C+O stars and the resultant type Ic supernova formation rate from the above paths using an initial mass function $\Psi(m) \propto m^{-2.7}$ and a q -distribution $\Phi(q) = 2(1+q)^{-2}$. We estimate the probability S that the binary survives after the spiral-in without merging to be $S \approx 0.25$ (path A) and 0.5 (paths B and C) from the relative range of initial separation. The fraction of binaries that evolve conservatively as a function of primary mass is taken from ref. 18. We then obtain the relative rates of type Ic with respect to type II supernovae as ~ 0.015 (path A), 0.025 (B) and 0.065 (C) for a close binary fraction of 0.5. (Note that path C is the most important channel, mainly due to the shape of the initial mass function.) In total, the expected type Ic formation frequency is ~ 0.1 times that of type II supernovae. This is probably consistent with the observed relative ratio between Ic and II if type Ib and Ic rates are comparable (ref. 19; but also see ref. 20). A more detailed estimate using Monte Carlo simulations is in progress.

The above C+O star model for type Ic supernovae can be substantiated from the comparison between the hydrodynamic model of explosion and the observed light curve and spectra of SN1994I and type Ic supernovae. We adopt the evolutionary model that a $15M_\odot$ main-sequence star becomes a helium star of $M_\alpha = 4.0M_\odot$, which subsequently loses its helium envelope to become a bare C+O star of $M_{C+O} = 2.1M_\odot$ (ref. 16). (This C+O star can form through path C if both star 1 and star 2 are more massive than $\sim 9M_\odot$.) We construct the pre-supernova C+O star model by replacing the $1.9M_\odot$ helium envelope of the pre-supernova helium star with a $0.14M_\odot$ C+O envelope. Elemental mass fractions near the surface are $X_{\text{Mg}} = 0.01$, $X_{\text{Ne}} = 0.02$, $X_{\text{O}} = 0.43$, $X_{\text{C}} = 0.45$, $X_{\text{He}} = 0.09$ and heavier elements with solar abundances. This C+O star model is denoted CO21.

This C+O star undergoes collapse due to photodisintegration of Fe. To simulate the subsequent explosion by generating a strong shock wave, we assume thermal energy deposition at the mass cut that divides a neutron star of $1.36M_\odot$ and ejecta. The resulting ejecta has a mass $M_{\text{ej}} = 0.88M_\odot$ and kinetic energy of explosion $E = 1 \times 10^{51}$ erg. It is noticeable that the ejecta mass is smaller than the Chandrasekhar-mass white-dwarf model.

Nucleosynthesis yields are nearly the same as in the explosion of the $4M_\odot$ helium star⁸. Produced heavy elements include $0.07M_\odot$ ^{56}Ni , $0.0033M_\odot$ Ca, $0.023M_\odot$ S, $0.071M_\odot$ Si, $0.050M_\odot$ Mg, $0.046M_\odot$ Ne, $0.50M_\odot$ O and $0.10M_\odot$ C. Unlike

the helium star, development of Rayleigh–Taylor instabilities and material mixing are not seen, because the density jump across the O/Si interface is not significant in C+O stars.

Because the C+O star is compact, with a radius of $\sim 0.2R_\odot$, the light curve is not due to shock heating but is powered by radioactive decay chain of $^{56}\text{Ni} \rightarrow ^{56}\text{Co} \rightarrow ^{56}\text{Fe}$. The optical light curve is calculated as in ref. 8, assuming 2 values of optical capacity $\kappa = 0.1$ and $0.2 \text{ cm}^2 \text{ g}^{-1}$ in view of uncertainties in line opacities at a steep velocity gradient. No mixing of ^{56}Ni is assumed.

Figure 1 compares the theoretical bolometric light curve of CO21 with other explosion models, as well as the observed bolometric light curve of SN1994I and other supernovae (see Figure legend for details). The light curve of CO21 is in good agreement with the observed curve of SN1994I. Compared with other theoretical models and other types of supernovae (except for the peculiar type Ia SN1991bg), the light-curve width is significantly narrower and the decline of the tail is much faster for CO21 and SN1994I. This is because radiative diffusion and γ -ray escape are faster because of the smaller ejecta mass and higher expansion velocities. The agreement at $\sim 5\text{--}10$ d after maximum can be improved if a realistic opacity (P. Höflich, personal communication) is used.

The ejecta mass is well constrained from the decline rate of the light-curve tail, which does not depend on κ but depends on the γ -ray energy deposition rate shown by the dashed line in Fig. 1. Because the γ -ray optical depth and thus the slope of the tail depends on $M_{\text{ej}}/E^{1/2}$, the favoured ejecta mass scales as $M_{\text{ej}} = 0.88M_\odot (E/10^{51} \text{ erg s}^{-1})^{1/2}$ (ref. 21). Thus unless $E > 2.5 \times 10^{51} \text{ erg s}^{-1}$, the ejecta mass is well below the Chandrasekhar mass, which strongly favours the C+O star progenitor.

The mass of synthesized ^{56}Ni may be inferred from the luminosity. For SN1994I (Fig. 1) we adopted 7.7 Mpc for the distance to NGC5194 (ref. 22) and an extinction of $A_V = 1.0$ mag. With these assumptions, the maximum bolometric luminosity of SN1994I is estimated to be $L_{\text{bol}} \approx 1.8 \times 10^{42} \text{ erg s}^{-1}$ (B. Schmidt and R. Kirshner, personal communication), which corresponds to $\sim 0.07M_\odot$ ^{56}Ni . A greater distance and larger extinction would imply a larger ^{56}Ni mass.

A search for a weak hydrogen feature in type Ic supernovae²³ would help to distinguish between the evolutionary paths. For path A, a small amount of hydrogen could be ablated from a low-mass main-sequence companion by the explosion. For paths B and C where the companion is either a white dwarf or a neutron star, hydrogen would be absent. The outer C+O layers contain a small amount of helium, which may be consistent with the observed weak He features⁶.

Because of its small mass and high expansion velocities, the ejecta will soon enter the supernebula phase where spectra are dominated by emission lines¹. These C+O stars contain significant amount of Si, S, Ca and ^{56}Ni . Even without mixing, these heavy elements expand at high velocities up to $\sim 10,000 \text{ km s}^{-1}$ for CO21, leading to broad emission lines. Oxygen lines should be even broader. Also, the column density of the ejecta will decrease so rapidly that X-ray signatures of a possible pulsar could be observed with the X-ray astronomical satellite ASCA if its luminosity is similar to the Crab nebula²⁴.

The explosion of a C+O star is likely to leave behind a neutron star. Unless the kick velocity of this neutron star is too high, the binary system should survive because of the small ejecta mass. The companion of the neutron star would be a low-mass main-sequence star, a C+O white dwarf, or another neutron star in a short and eccentric orbit.

The model described here should account for the presence of circumstellar matter inferred from radio observations of SN1994I (ref. 25). In the C+O star model, such matter is expected to form from a remnant of a helium-rich common envelope as well as from stellar-wind material from the C+O star. Hot plasma formed from collision between the ejecta and circumstellar matter would therefore be rich in heavy elements,

mostly C+O, which may affect X-ray emission and cooling properties. □

Received 2 May; accepted 26 July 1994.

- Branch, D., Nomoto, K. & Filippenko, A. V. *Comments Astrophys.* **15**, 221–237 (1991).
- Wheeler, J. C. & Harkness, R. P. *Rep. Prog. Phys.* **53**, 1467–1557 (1990).
- Filippenko, A. V., Porter, A. C. & Sargent, W. L. W. *Astr. J.* **100**, 1575–1587 (1990).
- Wheeler, J. C. *et al. Astrophys. J.* **313**, L69–L73 (1987).
- Schmidt, B., Challis, P. & Kirshner, R. *IAU Circ. No.* 5966 (1994).
- Clocchiatti, A., Brotherton, M., Harkness, R. P. & Wheeler, J. C. *IAU Circ. No.* 5972 (1994).
- Woosley, S. E., Langer, N. & Weaver, T. A. *Astrophys. J.* **411**, 823–839 (1993).
- Shigeyama, T., Nomoto, K., Tsujimoto, T. & Hashimoto, M. *Astrophys. J.* **361**, L23–L27 (1990).
- Nomoto, K., Filippenko, A. V. & Shigeyama, T. *Astr. Astrophys.* **240**, L1–L4 (1990).
- Lucy, L. B. *Astrophys. J.* **383**, 308–313 (1991).
- Swartz, D. A., Filippenko, A. V., Nomoto, K. & Wheeler, J. C. *Astrophys. J.* **411**, 313–322 (1993).
- Yamaoka, H., Shigeyama, T. & Nomoto, K. *Astr. Astrophys.* **267**, 433–438 (1993).
- Wheeler, J. C. & Swartz, D. A. in *Evolution of Massive Stars: Confrontation Between Theory and Observation* (eds Vanbeveren, D. *et al.*) 425–437 (Vrije Univ., Brussels, 1994).
- Bhattacharya, D. & van den Heuvel, E. P. J. *Phys. Rep.* **203**, 1–124 (1991).
- Habets, G. M. H. J. *Astr. Astrophys.* **187**, 209–230 (1986).
- Nomoto, K. & Hashimoto, M. *Phys. Rep.* **163**, 13–36 (1988).
- Van den Heuvel, E. P. J. in *Interacting Binaries* (eds Nussbaumer, H. & Orr, A.) 263–474 (Springer, Berlin, 1994).
- Polis, O. R., Côté, J., Waters, L. B. F. M. & Heise, J. *Astr. Astrophys.* **241**, 419–438 (1991).
- Van den Bergh, S. A. *Rev. Astr. Astrophys.* **29**, 363–407 (1991).
- Müller, R. A. *et al. Astrophys. J.* **384**, L9–L13 (1992).
- Swartz, D. A. & Wheeler, J. C. *Astrophys. J.* **379**, L13–L16 (1991).
- de Vaucouleurs, G. *Astrophys. J.* **227**, 729–755 (1979).
- Jeffery, D. J., Branch, D., Filippenko, A. V. & Nomoto, K. *Astrophys. J.* **377**, L89–L92 (1991).
- Shigeyama, T. *et al. Astrophys. J.* **420**, 341–347 (1994).
- Rupen, M. P. *et al. IAU Circ. No.* 5963 (1994).
- Yamaoka, H., Nomoto, K., Shigeyama, T. & Thielemann, F.-K. *Astrophys. J.* **393**, L55–L58 (1992).
- Nomoto, K., Yamaoka, H., Shigeyama, T. & Iwamoto, K. in *Supernovae and Supernova Remnants* (eds McCray, R. & Wang, Z.) (IAU colloq. No. 145, Kluwer, Dordrecht, in the press).
- Panagia, N. in *High Energy Phenomena Around Collapsed Stars* (ed. Pacini, F.) 33–49 (Reidel, Dordrecht, 1987).
- Leibundgut, B. *et al. Astrophys. J.* **371**, L23–L26 (1990).
- Leibundgut, B. *et al. Astr. J.* **105**, 301–313 (1993).

ACKNOWLEDGEMENTS. We thank K. Weiler, T. Kato and B. Schmidt for information on SN1994I. This work has been supported in part by the Ministry of Education, Science and Culture in Japan.

Coma formation driven by carbon monoxide release from comet Schwassmann–Wachmann 1

Matthew C. Senay & David Jewitt

Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

DISTANT comets are sometimes observed to undergo outbursts of activity that generate a surrounding coma, but the cause of this activity is not known^{1,2}. Whereas such outbursts in near-Sun comets are driven by the sublimation of water ice³, distant comets are too cold for this process to operate. The most plausible mechanisms involve the release of trapped gases from ice heated by an exothermic phase transition from an amorphous to a crystalline state^{4–6}, or the sublimation of very volatile ices such as molecular nitrogen and carbon monoxide^{7–9}. Here we report the detection of emission from carbon monoxide at submillimetre wavelengths from a distant comet, the periodic comet Schwassmann–Wachmann 1. The inferred rate of CO production is sufficient to generate the observed coma. These results provide the first direct evidence that sublimation of volatiles can drive the activity of distant comets.

Periodic comet Schwassmann–Wachmann 1 (SW1) lies beyond Jupiter in a nearly circular orbit with a semimajor axis of 6.03 AU and a period of 14.8 yr. Its perihelion and aphelion are at 5.77 and 6.28 AU, respectively. The comet is famous both for its episodic outbursts of dust^{10,11} and for its transient optical

emission lines of singly ionized carbon monoxide^{12,13}. Optical photometry shows that SW1 is active throughout its orbit¹⁴ and displays a persistent dust coma when observed with modern imaging detectors (Fig. 1).

Carbon monoxide $J(2-1)$ emission was detected from SW1 with the 15-m James Clerk Maxwell Telescope (JCMT) atop Mauna Kea, Hawaii on 22 October, 11 and 12 November 1993 UT (ref. 15). On all three dates the line appeared near the anticipated geocentric velocity of the comet (Fig. 2 and Table 1) but blue-shifted by ~ 0.4 km s⁻¹. Astrometry with the University of Hawaii 2.2-m telescope showed SW1 to be within 2 arcsec of its ephemeris position. The pointing accuracy of the JCMT is better than 2 arcsec (r.m.s.) and the tracking better than 1 arcsec over periods of ~ 1 h. At 230 GHz, the beam diameter is ~ 21 arcsec full width at half maximum¹⁶ (FWHM), corresponding to 9×10^4 km at SW1 on 22 October (see Fig. 1).

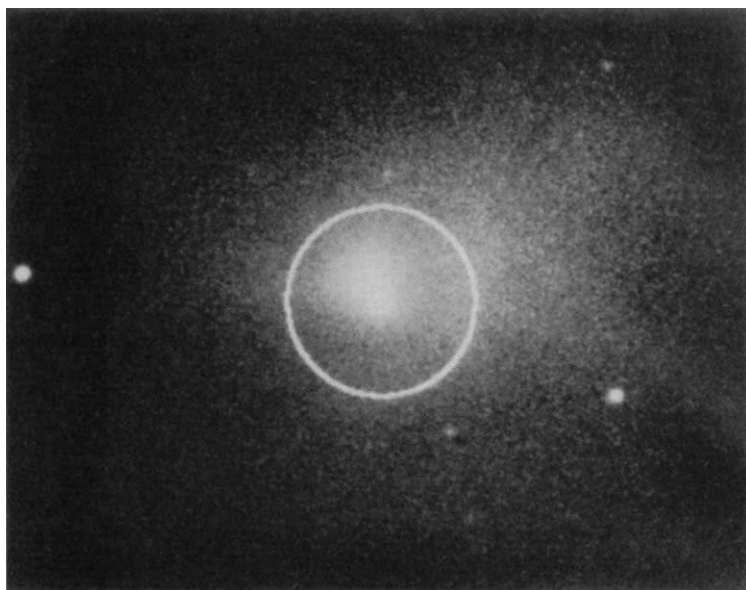
The blue shift of the CO line on each date is evidence for sublimation of CO near the subsolar point of the illuminated hemisphere, which produces a jet of gas and dust approximately towards the Sun. The velocity of the jet has a component projected along our line of sight equal to the observed blue shift of 0.4 km s⁻¹. The exact orientation of the jet cannot be determined from our observations. We adopt a velocity component of 0.28 km s⁻¹ normal to our line of sight as a reasonable estimate. The normal component of the jet velocity is used below to calculate the beam-crossing time for the CO molecules. Our value of 0.4 km s⁻¹ agrees well with the 0.5 km s⁻¹ blue-shifted CO $J(2-1)$ emission reported by Lellouch *et al.*¹⁷.

The observed line width for 22 October (our longest integration) is 0.7 ± 0.1 km s⁻¹ (FWHM). However, the A2 receiver used for our observations has a large (≥ 0.45 km s⁻¹) and uncertain instrumental width, due to instabilities in the receiver lock loop¹⁸. The JCMT data therefore provide only a crude upper limit to the intrinsic width of the CO line. A subsequent observation with the 10.4-m Caltech Submillimeter Observatory (CSO) on 13 May 1994 UT showed a deconvolved line of FWHM 0.20 ± 0.06 km s⁻¹. The half-width of 0.10 ± 0.03 km s⁻¹ provides a measure of the thermal velocity of the escaping CO. This is compatible with the 0.12 km s⁻¹ thermal velocity expected for freely subliming CO at 6 AU. This half-width is of the same order as the dust coma expansion velocity of 0.2 km s⁻¹ measured during outburst by Jewitt¹⁴, and the Bobrovnikoff expansion velocity of 0.18 km s⁻¹ (ref. 10) appropriate for the heliocentric distance at the time of observation (Table 1).

Production rates of CO (dM/dt) derived from our detections depend on the physical conditions prevailing in the coma. As we have observed only the $J(2-1)$ line, the accuracy of our estimate for dM/dt is limited by our lack of knowledge of the CO rotational level populations. Both fluorescence and thermal excitation are potentially important^{19,20}. The aperture-crossing time for gas at 0.28 km s⁻¹ is $t_a \approx 1.6 \times 10^5$ s, whereas the time to reach fluorescence equilibrium at 6 AU is $t_f \approx 10^6$ s (ref. 21). Therefore, we expect that the true rate of CO production is between the fluorescence and thermal limits. An additional uncertainty results from the possible release of CO from coma grains, as was reported in comet Halley²².

To bracket the likely production rates, we consider both fluorescence and thermal excitation of CO when applied to the line area of 0.08 ± 0.01 K km s⁻¹ (Table 1: 22 October). Fluorescence equilibrium is established when the rate of photoexcitation of the first vibrational level of CO by 4.7- μ m wavelength infrared photons balances the rate of spontaneous decays to lower-lying levels of the CO molecule. Under these conditions, the fractional population of each rotational level in the ground vibrational state is constant. For CO in SW1, $\sim 28\%$ of the molecules are in the $J=2$ rotational level of the ground vibrational state if fluorescence equilibrium prevails. The corresponding (temperature-independent) lower limit to the CO production rate is 1,500 kg s⁻¹ (3.1×10^{28} molecules per second).

FIG. 1 Charge-coupled device image of comet Schwassmann–Wachmann 1, taken with the University of Hawaii 2.2-m telescope on 12 November 1993 at 12:20 UT. This 100-s integration through a broadband red filter (central wavelength, 6,475 Å; FWHM, 900 Å) shows the extent and distribution of the dust coma. The circle denotes the 21-arcsec diameter JCMT beam. North is to the top, east to the right in this Figure.



In thermal equilibrium, the Boltzmann equation allows us to calculate the fractional level populations for a given kinetic temperature. For our purposes, the rates of collisionally induced transitions into and out of any particular level are then equal. The temperature of a freely sublimating CO surface at 6 AU is ~ 26 K. We consider CO gas at 20 and 80 K to illustrate the sensitivity of the mass production rate to the adopted temperature. At 20 K, we find a fractional population in the $J=2$ level of the ground vibrational state of 29% and a CO production rate of $1,400 \text{ kg s}^{-1}$ (3.0×10^{28} molecules s^{-1}). At 80 K, the corresponding quantities are 14% and $2,900 \text{ kg s}^{-1}$ (6.2×10^{28} molecules s^{-1}), respectively.

As a compromise between these extremes, we adopt a CO production rate of $\sim 2,000 \text{ kg s}^{-1}$ (4.3×10^{28} molecules s^{-1}), while recognizing that it is uncertain by a factor of the order of 2. This value for dM/dt agrees with the limit $dM/dt > 470 \text{ kg s}^{-1}$ (10^{28} molecules s^{-1}), inferred²³ from observations of the CO^+ ion tail of SW1. On 12 November (Fig. 1), the integrated red magnitude of the dust coma within a circle 10 arcsec in radius centred on the comet was 15.02 ± 0.05 . From comparison with published photometry¹⁴ we conclude that SW1 was in a quiescent activity state, with no evidence for a recent outburst.

A steady dust production rate of $\sim 10 \text{ kg s}^{-1}$ in micrometre-sized dust particles was deduced for SW1 from broadband photometry¹⁴, whereas independent models of the dust coma isophotes give $300\text{--}900 \text{ kg s}^{-1}$ (ref. 24). The dust production could be even greater if large (millimetre- to centimetre-sized) particles are abundant. Such particles are suggested by the detection of a dust trail by IRAS²⁵. But our CO production rate exceeds even the larger estimates of dust production, showing that CO can drive the optically dominant dust coma.

The asymmetric dust coma (Fig. 1) indicates local sites of CO sublimation rather than uniform activity. We estimate the percentage of active surface area on SW1 by considering the

range of possible values for the latent heat of sublimation L_{CO} of CO under different circumstances of association with water ice. Pure CO has $L_{\text{CO}} = 3 \times 10^5 \text{ J kg}^{-1}$ (ref. 26). Carbon monoxide adsorbed on water ice has $L_{\text{CO}} = (3.9\text{--}5.2) \times 10^5 \text{ J kg}^{-1}$, whereas larger values of $(1.1\text{--}1.7) \times 10^6 \text{ J kg}^{-1}$ apply to CO trapped in water ice²⁷. The total sublimating area (in m^2) is given by

$$a_s = \frac{dM/dt R^2 L_{\text{CO}}}{F_{\text{Sun}}(1-A)}$$

where $dM/dt = 2,000 \text{ kg s}^{-1}$ is the observed CO production rate, $R = 6.08 \text{ AU}$ is the heliocentric distance, $F_{\text{Sun}} = 1,360 \text{ W m}^{-2}$ is the solar flux at 1 AU, and $A = 0.15$ is the adopted Bond albedo²⁸. Using the range for L_{CO} given above, we find that a_s varies from a minimum of $1.9 \times 10^7 \text{ m}^2$ (equivalent circular radius 2.5 km) for pure CO to a maximum of $1.1 \times 10^8 \text{ m}^2$ (equivalent circular radius 5.9 km) for CO trapped within water ice. These correspond to active surface area percentages of 0.33% and 2.2%, respectively, assuming a spherical nucleus of 20 km radius²⁸ (smaller radii have been recently suggested²⁹). These estimates are very similar to the fractional active areas measured for other comets³⁰.

We derive a crude estimate for the maximum timescale of surface vent activity, before self-shadowing or collapse of the vent walls reduces sublimation³¹. This time is reached when the depth of the vent approximately equals its radius. For a cylindrical vent, the self-shadowing time is

$$\tau \approx \frac{\pi \rho r_s^3}{dM/dt}$$

where r_s is the vent radius, ρ is the density of the cometary dust-ice mixture and dM/dt is the total mass loss rate. Substituting $\rho = 10^3 \text{ kg m}^{-3}$, $r_s = 5.9 \text{ km}$ (the largest possible value for CO

TABLE 1 Observational circumstances and results

Date (UT, 1993)	R (AU)	Δ (AU)	α (degrees)	t (s)	V_g (km s^{-1})	V_o (km s^{-1})	$T_A \cdot dv$ (K km s^{-1})
Oct 22.67	6.075	5.905	−9.39	9,180	−27.21	-27.62 ± 0.05	0.08 ± 0.01
Nov 11.53	6.077	5.581	−8.52	3,000	−24.85	-25.25 ± 0.06	0.09 ± 0.01
Nov 12.53	6.077	5.589	−8.45	1,800	−24.64	-25.03 ± 0.04	0.05 ± 0.01

Abbreviations used: R , Sun–comet distance ($1 \text{ AU} = 1.5 \times 10^{11} \text{ m}$); Δ , Earth–comet distance; α , Sun–comet–Earth angle; t , integration time; V_g , expected geocentric radial velocity; V_o , observed geocentric radial velocity; $T_A \cdot dv$, line area (T_A is antenna temperature, dv is incremental line width in km s^{-1}).

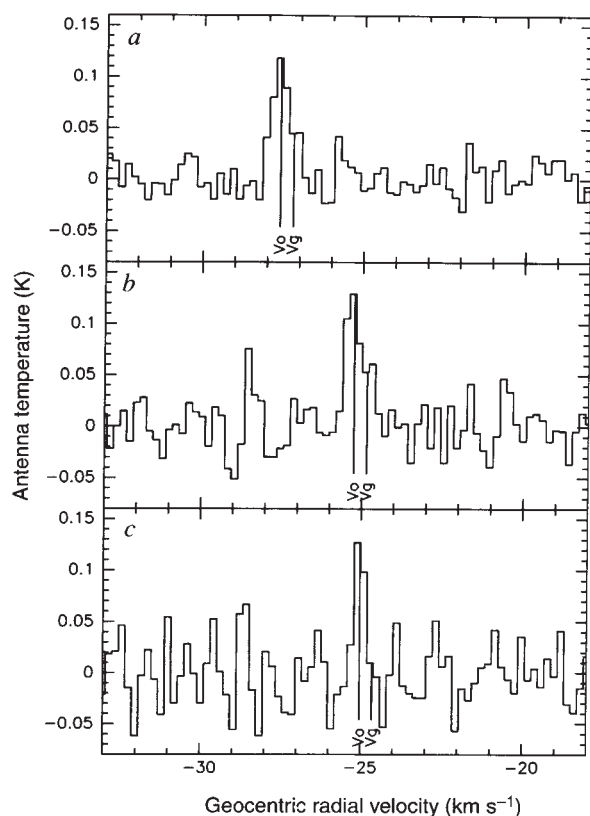


FIG. 2 JCMT observations of CO $J(2-1)$ emission from Schwassmann-Wachmann 1 (SW1). The CO $J(2-1)$ rotational transition at 230 GHz was observed using the A2 receiver (receiver temperature $T_{rx} \approx 95$ K, single sideband), and the digital autocorrelation spectrometer (125 MHz total bandwidth at a resolution of 95 kHz per channel). Spectral line widths in the cold comae of distant comets are expected to be < 1 km s $^{-1}$ (0.77 MHz at 230 GHz). For an observed system temperature $T_{sys} \approx 350$ K, our 3σ antenna temperature, $T_A^*(r.m.s.)$, detection limit in 1 h of position-switching integration at 1 km s $^{-1}$ resolution is ~ 0.046 K, equivalent to a flux density of ~ 1.1 Jy, assuming an antenna efficiency of 24.8 Jy K $^{-1}$ (ref. 16). a, CO $J(2-1)$ emission from SW1 recorded on 22 October 1993. This and the following spectra have been smoothed to a resolution of ~ 0.2 km s $^{-1}$. The peak antenna temperature is ~ 0.1 K and the FWHM $\approx 0.7 \pm 0.1$ km s $^{-1}$ for this strongest detection, achieved in 9,180 s of position-switching integration time. This translates into a flux density ~ 2.8 Jy. The expected geocentric radial velocities of the CO line for each observation date are indicated by the vertical lines labelled V_g in the figure panels. The observed CO line radial velocities for each date are indicated by vertical lines labelled V_o . b, 3,000-s integration showing CO $J(2-1)$ emission from SW1 on 1 November 1993. c, 1,800-s integration showing CO $J(2-1)$ emission from SW1 on 12 November 1993.

trapped in water ice) and $dM/dt \approx 2,000$ kg s $^{-1}$, we find that $\tau \approx 10^4$ yr, or 700 orbits. For a pure CO patch $\tau \approx 800$ yr, or 60 orbits. As on the nucleus of comet Halley³², there are probably several, smaller vents each with a correspondingly shorter lifetime. For CO trapped in water ice, a 100-m diameter vent has a lifetime of ~ 100 yr, still considerably longer than the ~ 15 yr orbital period. It thus appears plausible that CO, outgassing from localized vents, drives the observed persistent dust coma.

In summary, we have detected CO submillimetre emission from a comet. The CO production rate we infer is sufficient to produce the observed amounts of dust and CO $^+$, the two outstanding features of optical activity in SW1. Although our observations do not eliminate other mechanisms, and do not exclude activity caused by other very volatile gases (such as N $_2$), this detection provides strong evidence that CO is an important driver of activity in distant comets. $\square$

Received 18 March; accepted 26 July 1994.

1. Roemer, E. *Publ. astr. Soc. Pacif.* **74**, 351–365 (1962).
2. Sekanina, Z., Larson, S. M., Hainaut, O., Smette, A. & West, R. M. *Astr. Astrophys.* **263**, 367–386 (1992).
3. Whipple, F. L. *Astrophys. J.* **111**, 375–394 (1950).
4. Klinger, J. *Science* **209**, 271–272 (1980).
5. Smoluchowski, R. *Astrophys. J.* **244**, L31–L34 (1981).
6. Bar-Nun, A., Dror, J., Kochavi, E. & Laufer, D. *Phys. Rev.* **B35**, 2427–2435 (1987).
7. Cowan, J. J. & A'Hearn, M. F. *Icarus* **50**, 53–62 (1982).
8. Fanale, F. P. & Salvail, J. R. *Icarus* **84**, 403–413 (1990).
9. Crovisier, J. in *Workshop on the Activity of Distant Comets* (eds Huebner, W. F., Keller, H. U., Jewitt, D., Klinger, J. & West R.) 153–160 (Southwest Research Institute, San Antonio, 1992).
10. Whipple, F. H. *Astr. J.* **85**, 305–313 (1980).
11. Hughes, D. W. in *Comets in the Post-Halley Era* (eds Newburn, R. L. Jr, Neugebauer, M. & Rahe, J.) 825–851 (Kluwer, Dordrecht, 1991).
12. Larson, S. M. *Astrophys. J.* **238**, L47–L48 (1980).
13. Cochran, A. L. & Cochran, W. D. *Icarus* **90**, 172–175 (1991).
14. Jewitt, D. *Astrophys. J.* **351**, 277–286 (1990).
15. Senay, M. C. & Jewitt, D. *IAU Circ.* No. 5929 (1994).
16. Matthews, H. E. in *The James Clerk Maxwell Telescope: A Guide for the Prospective User* (ed. Matthews, H. E.) 1–25 (Joint Astronomy Center, Hilo, 1993).
17. Lellouch, E. et al. *IAU Circ.* No. 5994 (1994).
18. Matthews, H. E. *JCMT Newsletter* No. 2, 5–7 (Royal Observatory, Edinburgh, 1994).
19. Chin, G. & Weaver, H. A. *Astrophys. J.* **285**, 858–869 (1984).
20. Bockelee-Morvan, D. & Crovisier, J. *Astr. Astrophys.* **151**, 90–100 (1985).
21. Crovisier, J. & Le Boulou, J. *Astr. Astrophys.* **123**, 61–66 (1983).
22. Eberhardt, P. et al. *Astr. Astrophys.* **187**, 481–484 (1994).
23. Jockers, K., Bonev, T., Ivanova, V. & Rauer, H. *Astr. Astrophys.* **260**, 455–464 (1992).
24. Fulle, M. *Nature* **359**, 42–44 (1992).
25. Sykes, M. V. & Walker, R. G. *Icarus* **95**, 180–210 (1992).
26. Brown, G. N. & Ziegler, W. T. *Adv. cryogen. Engng* **25**, 662–670 (1979).
27. Schmitt, B. in *Interrelations Between Physics and Dynamics for Minor Bodies in the Solar System* (eds Benest, D. & Froeschle, C.) 265–307 (Editions Frontieres, Gif-sur-Yvette, 1992).
28. Cruikshank, D. P. & Brown, R. H. *Icarus* **56**, 377–380 (1983).
29. Meech, K. J., Belton, M. J. S., Mueller, B. E. A., Dickson, M. W. & Li, H. R. *Astr. J.* **106**, 1222–1236 (1993).
30. Jewitt, D. in *Comets in the Post-Halley Era* (eds Newburn, R. L. Jr, Neugebauer, M. & Rahe, J.) 19–65 (Kluwer, Dordrecht, 1991).
31. Sekanina, Z. *Astr. J.* **100**, 1293–1314 (1990).
32. Keller, H. U. in *Physics and Chemistry of Comets* (ed. Huebner, W. F.) 13–68 (Springer, Berlin, 1990).

ACKNOWLEDGEMENTS. We thank C. Purton, H. Matthews and R. Tilanus for scientific support, and A. Hatakayama, E. Lundin, R. Luthe and J. Cox for observing assistance, at JCMT. We also thank A. Evans and D. Sanders for observing time, J. Deane for assistance at the CSO, and S. Ridgway for obtaining the image used in Fig. 1. N. Ladd provided advice on data reduction. M.C.S. and D.J. are visiting astronomers at the James Clerk Maxwell Telescope, operated by the Observatories on behalf of the UK SERC, the Netherlands Organization for Scientific Research, and the Canadian NRC. This work was supported by NASA (D.J.) and the ARCS Foundation (M.C.S.).

Bubbling in vertically vibrated granular materials

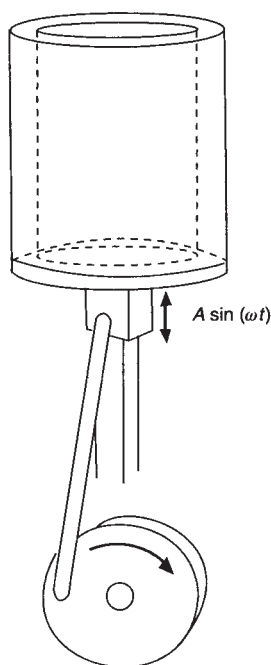
H. K. Pak & P. R. Behringer

Department of Physics and Center for Nonlinear and Complex Systems, Duke University, Durham, North Carolina 27708-0305, USA

GRANULAR materials show both fluid-like and solid-like behaviour. Under weak shear they deform plastically; under high shear they flow. These materials exhibit other unusual kinds of behaviour, including segregation¹, density waves², convection³ and anomalous sound propagation⁴. Their dynamical properties are important in many industrial applications^{5–7}. In particular, the shaking of granular materials is used to mix, segregate and transport them. Vertically shaken granular materials undergo a transition to a convective state^{6–14}. Here we describe experiments which show that such convective motion can involve bubbling—the formation and upward motion of voids. The presence of gas between the grains is essential for bubbling to occur, and the instability shows characteristics of a Hopf bifurcation such as is seen at the onset of chaos. This bubbling behaviour may be analogous to that observed in fluidized beds^{15,16}, and might be expected to occur when soils are fluidized during earthquakes.

Granular materials lying on a vertically vibrating surface are a topic of recent interest^{6–14} for a number of reasons. From the point of view of granular mechanics, this system incorporates many aspects of those materials, including the fluid-like and solid-like phases. The mechanisms leading to convection are still

FIG. 1 Sketch of the apparatus. The annular geometry eliminates the influence of one pair of lateral boundaries, giving a periodic boundary condition in the azimuthal direction. The cell is made of two concentric Plexiglas cylinders. The diameter of the inner cylinder is ~ 9.60 cm, and the radial gap size is 0.625 cm. In this experiment, vibration frequencies are in the range $0 < \omega/2\pi \leq 30$ Hz.



not well established; no theory predicts its general occurrence, in spite of work dating to the time of Faraday³. A continuing topic of debate is the role, if any, played by the surrounding gas. This system is also interesting from the point of view of nonlinear dynamics and chaos.

Assuming sinusoidal oscillations, $z = A \sin(\omega t)$, an important parameter is $\Gamma = a\omega^2/g$ (where g is the acceleration due to gravity, A the physical amplitude of the vibrating floor and ω the angular frequency of the vibration). Thus Γ measures the peak acceleration relative to gravity, and $\Gamma > 1$ is a necessary condition for sustained flow. Throughout the experiments reported here, Γ is varied by changing ω at fixed A , although several different values of A were used. As Γ increases above a critical value Γ_c , the free surface of the granular material becomes unstable. Typically, a single heap forms^{3,8-10}. When Γ

increases higher above Γ_c , there are two additional instabilities, depending on the height h of the layer compared to the particle diameter d : travelling surface waves¹² for large h/d , and sub-harmonic bifurcations^{13,14} for moderate h/d .

Here we describe experiments with rough sand ($d = 0.20$ mm) in a thin annulus of circumference $C = 32$ cm. The vibration axis is the cylindrical axis, and $h/d \gg 1$. (Typically, h is $3-4$ cm.) We find a novel bubbling effect, related to the presence of surrounding gas (air), which occurs for $\Gamma > \Gamma_b = 6.7 > \Gamma_w > \Gamma_c$. Surface waves occur for $\Gamma > \Gamma_w$. Above Γ_b , voids, reminiscent of bubbles in fluidized beds^{15,16}, are created near the bottom of the heap, move upwards and burst through the free surface. To our knowledge, this is the first report of bubbling in a shaken container of granular material. We note that as with the wavy states^{12,13}, some of the details, in particular Γ_b , will depend on d , h , A and other system parameters. Relevant dimensionless parameters in addition to Γ and h/d are A/d (see ref. 12) and C/d . But the importance of surrounding gas means that other quantities are likely to be relevant. For instance, air is more easily trapped in materials consisting of small grains, which is an effect depending on the gas viscosity.

The experiments are carried out using a mechanical shaker driven by a d.c. motor¹², as sketched in Fig. 1. The onset of convection and flow is found to occur at $\Gamma_c = 1.17 \pm 0.05$, in agreement with other work^{8,10,13}. Just above Γ_c , steady convective motion evolves in the form of two counter-rotating rolls leading to a single symmetrical heap.

Figure 2 shows a sequence of images characterizing the bubbles. The qualitative features in this figure are as follows. Bubbles are created near the bottom centre of the heap. As the bubbles move upwards, their mean size grows, but during each shaker cycle, bubbles are compressed and expanded. Bubbles near each other tend to coalesce and become one large bubble. As a bubble approaches the free surface of the system, it bursts out, creating a spray of sand (this may be related to the chaotic states reported previously^{7,9}). As Γ increases, the mean size of the bubbles increases, but the rate of bubbling is not very sensitive to Γ . Large bubbles can undergo another instability which exhibits a characteristic finger-like structure inside the bubble (Fig. 2d).

The onset of bubbling is related to a qualitative change in the flow. Just above Γ_c , the convective motion consists of two large

FIG. 2 Images of the bubbling of sand. Here $\Gamma = 8.79$, $\omega/2\pi = 22.85$ Hz, $A = 0.420$ cm and $h = 4.0$ cm. Times t are 0 (a), 0.17 s (b), 0.33 s (c) and 0.53 s (d). In d, there is a finger-like pattern inside the bubble. This feature always occurs just before or during the moment of bursting. The distance between fingers is approximately the size of the gap, suggesting that the size of the fingers is set by this length.



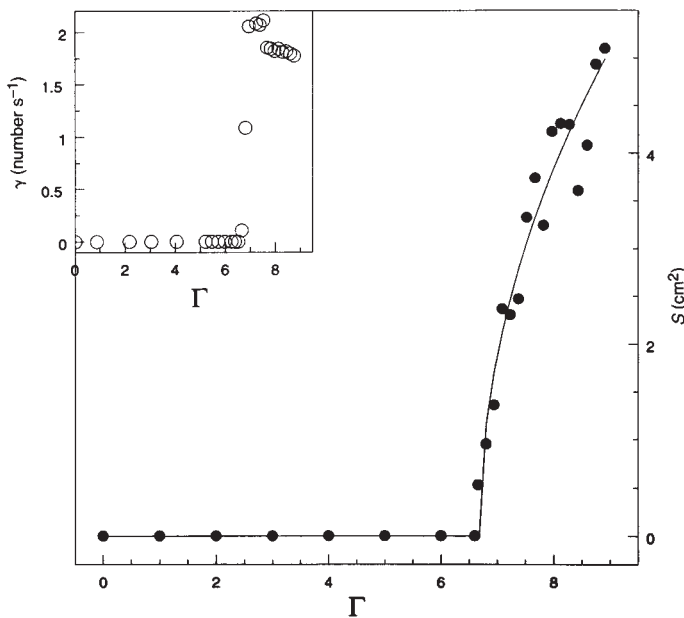


FIG. 3 Average bubble areas as a function of Γ . The solid circles indicate averages over several measurements at the same height and Γ . Inset, rate of bubble formation γ as a function of Γ . Here, $A = 0.543$ cm and $h = 3.0$ cm.

convective rolls with flow from the bottom of the heap towards the top. Just below Γ_b , however, the convective motion near the top reverses, so that sand flows down in the centre, where formerly it flowed up. At Γ_b the top of the heap starts to collapse, making a depression and two heaps. As Γ increases further, the distance between the two heaps becomes larger. At even larger Γ , bubbling effectively flattens the heaps. This is the situation in Fig. 2, where little if any of the original heap is visible.

The instability at Γ_b is a critical transition; that is, no hysteresis is apparent on cycling through the onset of bubbling. The size of the depression is a continuous reversible function of Γ . There is also no evident hysteresis in the rate of bubble generation or the size of the bubbles. The solid circles in Fig. 3 show the bubble area S (measured at a fixed height and averaged over many bubbles) versus Γ . The rate, γ , at which bubbles form is given in the inset. We find that γ rises rapidly with Γ , decreases slightly, and then is essentially constant. The data are reminiscent of a Hopf bifurcation in which the amplitude of oscillation, as measured here by the bubble size, grows as the square root of the bifurcation parameter, $\Gamma - \Gamma_b$, whereas the frequency of oscillation is essentially constant. In Fig. 3, a solid line is fitted by $S = a(\Gamma - \Gamma_b)^{1/2}$ with $a = 3.38$ cm and $\Gamma_b = 6.67$. This is an imperfect analogy, however, because the production of bubbles is not strictly a periodic function of time. However, the data suggest that the bubbles may be generated by a Hopf bifurcation in the presence of noise.

The effect of air is very important in the dynamics of bubble formation. To investigate the role of air, we varied the air pressure inside the shaker. At fixed Γ , the bubbling rate decreases with decreasing pressure and no bubbles form when all the air is evacuated from the shaker.

The nucleation of the bubbles is particularly interesting. Bubbles do not appear to form during the part of the cycle when the base and sand collide, that is, when air is forced into the granular material from the bottom. Rather, they appear to form during the part of the shaker cycle when the material leaves the shaker base and is in free flight. Sand near the bottom of the heap bulges downward, nucleating a bubble. This suggests that the bubbles form because compressed air trapped in the material can expand during the free-flight part of the cycle. An interesting

issue is the connection between bubbling and the downward convective flow which begins at the peak for Γ , just below Γ_b .

The formation of bubbles (or voids) is also observed in fluidized beds. In a fluidized bed, gas is injected from below into a column of granular material^{15,16}. The relation between the bubbling process seen here and that found in fluidized beds is an interesting topic for future investigation. □

Received 21 February; accepted 4 August 1994.

- Williams, J. C. *Powder Technol.* **15**, 245–251 (1976).
- Baxter, G. W., Behringer, R. P., Fagert, T. & Johnson, G. A. *Phys. Rev. Lett.* **62**, 2825–2828 (1989).
- Faraday, M. *Phil. Trans. R. Soc.* **121**, 299–340 (1831).
- Liu, C.-h. & Nagel, S. R. *Phys. Rev. Lett.* **68**, 2301–2304 (1992).
- Gutman, I. *Industrial Uses of Mechanical Vibration* (Business Books, London, 1968).
- Jaeger, H. M. & Nagel, S. R. *Science* **255**, 1523–1531 (1992).
- Evesque, P. *Contemp. Phys.* **33**, 245–261 (1992).
- Evesque, P. & Rajchenbach, J. *Phys. Rev. Lett.* **62**, 44–47 (1989).
- Laroche, D., Douady, S. & Fauve, S. *J. Phys., Paris* **50**, 699–706 (1989).
- Clément, E., Duran, J. & Rajchenbach, J. *Phys. Rev. Lett.* **69**, 1189–1192 (1992).
- Knight, J. B., Jaeger, H. M. & Nagel, S. R. *Phys. Rev. Lett.* **70**, 3728–3731 (1993).
- Pak, H. K. & Behringer, R. P. *Phys. Rev. Lett.* **71**, 1832–1835 (1993).
- Douady, S., Fauve, S. & Laroche, C. *Europhys. Lett.* **8**, 621–627 (1989).
- Melo, F., Umbanhowar, P. & Swinney, H. *Phys. Rev. Lett.* **72**, 172–175 (1994).
- Howard, J. R. *Fluidized Bed Technology, Principles and Applications* (Hilger, New York, 1989).
- Tsinontides, S. C. & Jackson, R. *Proc. Joint DOE/NSF Workshop on Flow of Particulates and Fluids* (eds Playnski, S. I. et al.) 254–286 (Pittsburgh Energy Technology Center, Pittsburgh, 1991).

ACKNOWLEDGEMENTS. This work was supported by the US NSF, the US Air Force Office of Research, and the Exxon Research and Engineering Co.

Persistent tangled vortex rings in generic excitable media

A. T. Winfree

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA

EXCITABLE media are exemplified by a range of living systems^{1–8}, such as mammalian heart muscle⁶ and its cells¹ and *Xenopus* eggs^{2,3}. They also occur in non-living systems such as the autocatalytic Belousov–Zhabotinsky reaction^{9–14}. In most of these systems, activity patterns, such as concentration waves, typically radiate as spiral waves from a vortex of excitation created by some non-uniform stimulus. In three-dimensional systems, the vortex is commonly a line, and these vortex lines can form linked and knotted rings which contract into compact, particle-like bundles^{9–30}. In most previous work these stable ‘organizing centres’ have been found to be symmetrical and can be classified topologically. Here I show through numerical studies of a generic excitable medium that the more general configuration of vortex lines is a turbulent tangle, which is robust against changes in the parameters of the system or perturbations to it. In view of their stability, I suggest that these turbulent tangles should be observable in any of the many known excitable media.

Excitable media sustain the propagation of a shock front with fixed speed and profile if, and only if, excitation is fast enough relative to recovery and large enough relative to the least triggering stimulus³¹. With only slight further increase in either parameter, rotors also become viable; this is why almost all excitable media exhibit spiral waves, given an appropriate initial stimulus. In this marginal case (the most studied, indeed exclusively so until a decade ago) the rotor is stationary and its spiral pivots rigidly. With a further adjustment of parameters, more typical behaviour emerges: the rotor exhibits spontaneous meander. Depending delicately on parameters, this spontaneous motion has various forms, some quasiperiodic and some not^{13,28,31–34}.

For convenience of experiment and image processing, most laboratory work still focuses on two-dimensional (2D) situations. In the few cases permitting 3D investigation, the 2D vortex becomes a vortex line, often a vortex ring: such examples include the Belousov-Zhabotinsky medium^{9,14}, the developing grex of the slime mould⁵ and heart muscle at the onset of ventricular fibrillation and sudden cardiac death^{35,36}. Vortex filaments in 3D differ from 2D vortices. The latter occur in only two forms: clockwise and anticlockwise rotors of fixed period. They cannot be created or extinguished except in such pairs or by perturbations reaching as far as the medium's boundary^{37,38}.

They can meander in an arbitrary and unchanging direction established by initial conditions and liable to redirection by delicate stimuli³¹⁻³⁴. In contrast, in 3D there is no distinction between clockwise and anticlockwise, and filaments typically close in rings which can spontaneously shrink and vanish to leave the medium again quiescent and excitable. They also have additional properties of 'curvature' and 'twist' which locally alter the rotation period and thus make meander more complex^{9,12,13,15,22,28,34}. 'Twist' is a helical rotation of concentration gradients along the length of a filament. Most importantly, in 3D there is a surprising diversity of topologically distinct ways

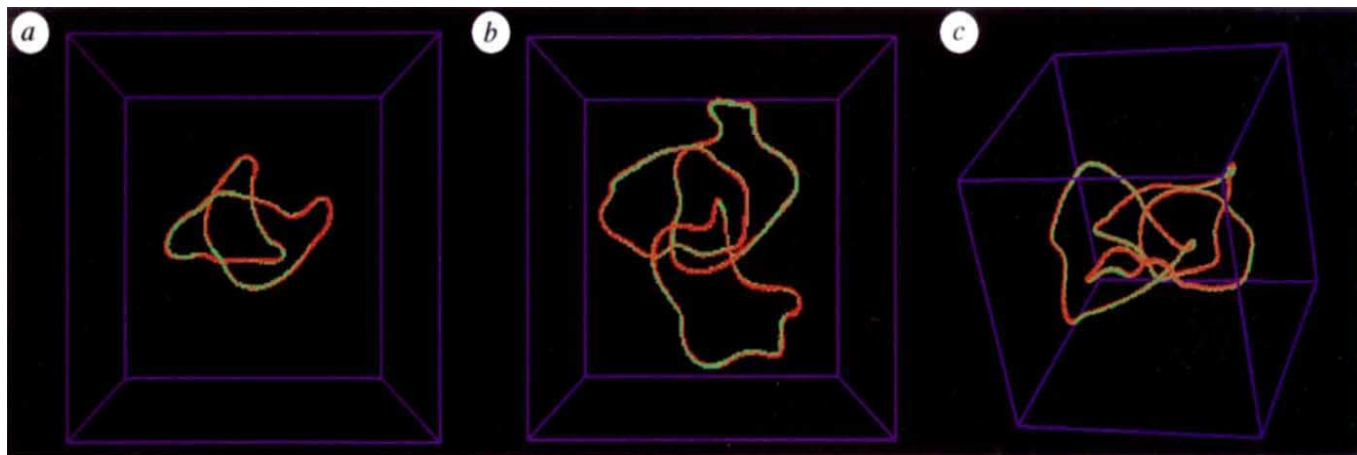
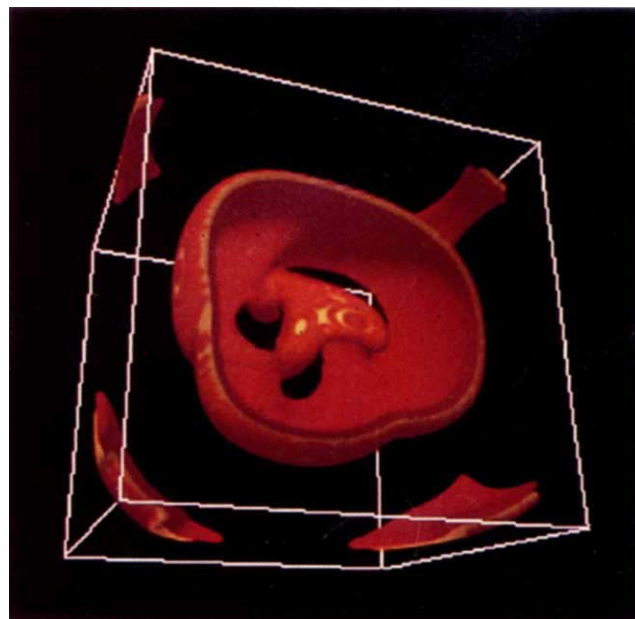


FIG. 1 The reaction-diffusion equations used in this figure are $\partial u/\partial t = (u - (u^3/3) - v)/\varepsilon + \nabla^2 u$ and $\partial v/\partial t = \varepsilon(u + \beta - (v/2))$ where u represents local membrane potential and v represents conductivity of membrane ionic channels involved in repolarization^{20,22,28,29}. At the unique parameters $\varepsilon = 0.3$ and $\beta = 0.7$ used in the past to maintain stable organizing centres^{20,29} this equation has a neutrally stable steady state; there are two alternative spiral solutions, with periods 11 and 17 time units, and in both the inner tip of the 2D spiral wave pursues a perfectly circular path²⁷. Here $\varepsilon = 0.2$ and $\beta = 0.9$, so that the inner tip of the unique 2D spiral wave pursues a cycloidal course across the medium, traversing the distance of one wavelength in ~ 6 rotation periods of 16 time units. In 3D this vigorous meander is reined in and redirected by the tendency of other arcs of the same filament to do the same in a different direction. These tangles were initialized in a grid of 151^3 points, 3.6 wavelengths along its edge (blue lines) with impermeable walls. Each gridpoint is coloured red to green by its instantaneous u value, revealing the twist

of u gradients along the filament induced by linkage and knotting^{17,19,22,24,26,28}. a, Single vortex ring knotted to prevent symmetrical collapse, the first-discovered stable organizing centre³⁰. It has now been noticed in numerical experiments with six different excitable media^{26,29,30}. But these were all confined to the marginal strip of parameter space where spirals pivot rigidly and perfect symmetry and periodicity prevail. This snapshot—from a more persuasive stereo video simulation available from the author for display on Silicon Graphics workstations—shows what becomes of it at a typical moment in a medium exhibiting the most violent meander. b, Three rings each linked once through the other two. c, Three rings neither linked nor knotted but inextricably entangled. This qualitatively supplements the decade-old classification of organizing centres, in which only linking and knotting were considered^{17,19,22,24}, and proves that those mathematical features are not essential for stability or persistence.

FIG. 2 The reaction-diffusion equations used in this Figure are $\partial u/\partial t = (u - (u^3/3) - v)/\varepsilon + \nabla^2 u$ and $\partial v/\partial t = \varepsilon(u + \beta - (v/2)) + \nabla^2 v$ where u and v might represent chemical concentrations, both diffusing equally. (In an ungelled Belousov-Zhabotinsky reaction, reagent v (ferroin), having higher molecular weight, might diffuse at 0.6 times the rate of u (bromous acid)^{12,34}. Such an adjustment has little effect on the outcome.) As in Fig. 1, ∇^2 is the laplacian operator, the second spatial derivative, representing molecular diffusion. Decreasing ε makes excitation more rapid and recovery slower. If $|\beta|$ is less than 0.7, the steady state is unstable towards a limit-cycle oscillation; larger values make excitation more difficult (higher threshold). Here $\varepsilon = 0.2$ and $\beta = 0.8$, so that the inner tip of the 2D spiral wave pursues a perfectly circular path of period 21 time units. At very slightly larger ε or β in this equal-diffusion medium, the rotor becomes non-viable and propagation fails. The grid (182^3 points) is 3.7 wavelengths along its edge (white lines), with impermeable walls. Instead of plotting gridpoints in the vortex core as in Fig. 1, I render by opaque colours chemical concentration u within a certain range in the wavefront, only between foreground and background planes. The inner edge of this wavefront (the complicated hole) is a trefoil-knotted vortex filament such as indicated by gridpoints in Fig. 1a, after 15 rotations at these parameters. In each plane perpendicular to any arc of the filament a spiral wave radiates outward. Such spiral slices collectively constitute the 3D wavefronts, which repeat periodically; only their innermost segments are visible here. (Graphics assistance from T. Elvins at San Diego Supercomputer Center (NetV) and M. Landis at the University of Arizona (Solitaire film recorder).)



that vortex rings can stably knot and link to become compact organizing centres^{17,19,22,24,28}. It would be of interest to clarify the classes of natural stimuli that instigate each of these many topologically distinct particles³⁷, here created by numerical prescription^{14,21,23,26,29}. Apart from the shrinking planar ring, these have so far been studied only by numerical solution of the known equations governing excitability in electrophysiological and chemical excitable media^{13,20–22,26,28–30}. After initial transients, each maintains a stable size and shape with its linked and/or knotted rings packed only ~ 1 vortex core-diameter apart (1 spiral wavelength/ π), radiating waves at an idiosyncratic period. Each slowly traverses the medium with a corkscrew-like rigid motion^{13,22,26,28,29}. All such activity can be erased by uniformly perturbing the medium outside the usual range of an excitation, whereupon it reverts to uniform quiescence³⁷. But left to spontaneous processes, these particles continue to emit excitations at a period somewhat shorter than the 2D vortex rotation, until, gliding at speeds on the order of 1% of the wave propagation speed, they run into a boundary or another such particle. Such compact organizing centres have not yet been sought in the laboratory, perhaps due in part to the suspicion that they are delicate with respect to anatomy and/or special parameter combinations required for their maintenance in numerical solutions of the pertinent reaction-diffusion equations.

The equations used to describe excitable media can be quite simple; here I use the FitzHugh-Nagumo equation long favoured for modelling electrical excitability in cell membranes^{20,22,28,29,31}. This equation describes any substance in which a spatially uniform steady state is stable against small perturbations, but a trans-threshold disturbance evokes a large, rapid but ephemeral local excitation. This propagates like a shock wave behind which the medium recovers to the same state of quiescent excitability. Propagation is mediated by diffusion: of electric potential in the electrophysiologist's cable equation in neuromuscular membrane systems, or of an excitator substance in chemical and biochemical systems with analogous kinetics. All past instances of compact organizing centres^{20,22,26,28–30} have arisen under two special conditions: (1) the quantitative values of excitability parameters are such that the medium is only marginally excitable and/or almost spontaneously oscillatory, so that the spiral wave pivots rigidly without the meander typical of fully excitable systems; and (2) only the excitator substance or potential diffuses. These non-generic conditions are not essential for the persistence of compact organizing centres, and such objects are not delicate. I show this by counter-example using computations in excitable media in which rotors meander (as in most electrophysiological and biochemical systems), or in which all reactants diffuse comparably (as in aqueous chemical solutions).

Meander has been observed previously only on twist-free uniformly curved or straight vortex filaments. I found that its pattern is affected if curvature and twist vary along the filament. The resulting asynchrony might be expected to disrupt delicately stable arrangements of adjacent vortex filaments: collisions and reconnections would then eventually undo all linkage, knotting or other entanglement, leaving only contracting rings that end in extinction and quiescence. To find out, I investigated three organizing centres known to be stable in marginally excitable media²⁹ by gradually or abruptly adjusting their parameters away from the atypical range in which rotors pivot rigidly. All immediately lose their former exquisite symmetries and periodicity: the filaments writhe and lash about but apparently also repel one another vigorously. All three arrangements of rings (Fig. 1) expanded somewhat and persisted through at least 40 rotation periods. No cumulative change of period or proportions was evident beyond the middle of this interval, whereas a ring of similar size and free of entanglements shrinks and vanishes in half that time.

In these three cases an odd number of rings were entangled. A case with an even number of rings was also tested: two rings (of same or different length and shape), each linked twice through the other. Asymmetry and aperiodicity arose similarly, and after slight initial expansion this pair also persisted without consistent change through the 40th rotation: but in contrast to the spontaneous loss of symmetry in Fig. 1, each ring, though always misshapen, remained perfectly symmetric through its centre of mass and even rebounded toward symmetry after one of the two filaments was 'plucked' in an asymmetric way by a deliberate numerical perturbation.

Because no shape ever repeated in the 2,644 frames of these four simulations, the 176 rotations reported here fall short of proving that such 'thermal agitation' will never jiggle the ostensibly persistent turbulent tangle into a shape from which it cannot recover without transmutation¹⁷ (for example, two linked rings connecting to form one ring) or even stochastic decay to quiescence (for example, the one ring shrinking and vanishing). But comparable persistence in laboratory analogues would suffice for observability.

In all known compact organizing centres only the fast-reacting local state variable was free to diffuse. To test whether this is necessary I made a trefoil-knotted ring (somewhat larger than proved stable) in a similar medium but with both variables freely diffusing. Its rotors do not meander; its circular twist-free rings shrink and vanish exactly as prescribed by the equal-diffusion theory of this case^{15,18}. Over the course of 15 rotations (long enough for a symmetric ring of comparable size to vanish) this knot shrank only slightly to a stable shape, thereafter maintained (Fig. 2). Thus, unlike Turing patterns, stable vortex rings do not require appropriate inequality of diffusion coefficients; they seem endemic to excitable media generally. Presumably the twist- and curvature-driven contraction of any arc is opposed ultimately by repulsion from skew segments on which it inevitably snags due to entanglement.

Mathematical analysis has been successful in describing the qualitative topological features of compact organizing centres^{17,19,22,24}, but never in predicting much about their persistence or quantitative dynamics^{13,20,22,26,29}. Existing approximations apply apparently only under the simplifying constraints that meander be suppressed and that filaments remain far apart and barely curved or twisted^{11,15,18,24}. Abandoning those restrictions to create more typical (and computable) cases revealed turbulent tangles that lack symmetry in space or periodicity in time, but persist topologically. They retain their initial entanglement while asymmetrically and aperiodically writhing about and emitting concentric shells of excitation at roughly the period of the 2D rotor. Although their motions still evade mathematical analysis, these robust solutions seem non-exceptional, so they should be observable in an adequate volume of any excitable medium. The dynamics of interaction among colliding particles, probably including topological transmutations¹⁷ and the scattering of glide directions among the survivors, would also merit investigation. $\square$

Received 31 May; accepted 8 August 1994.

1. Lipp, P. & Niggli, E. *Biophys. J.* **65**, 2272–2276 (1993).
2. Lechleiter, J., Girard, S. E., Peralta, E. G. & Clapham, D. E. *Science* **252**, 123–126 (1991).
3. Atri, A., Amundson, J., Clapham, D. & Sneyd, J. *Biophys. J.* **65**, 1727–1739 (1993).
4. Cornell-Bell, A., Finkbeiner, S. M., Cooper, M. S. & Smith, A. J. *Science* **247**, 470–473 (1990).
5. Steinbock, O., Siegert, F., Muller, S. C. & Weijer, C. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7332–7335 (1993).
6. Davidenko, J. M., Pertsov, A. M., Salomonsz, R., Baxter, W. & Jalife, J. *Nature* **355**, 349–351 (1992).
7. Gorelova, N. A. & Bures, J. *J. Neurobiol.* **14**, 353–363 (1983).
8. Delaney, K. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 669–673 (1994).
9. Winfree, A. T. *Scient. Am.* **230**(6), 82–95 (1974).
10. Pertsov, A. M., Vinson, M. & Muller, S. C. *Physica D* **63**, 233–240 (1993).
11. Pertsov, A. M., Aliev, R. R. & Krinsky, V. I. *Nature* **345**, 419–421 (1990).
12. Jahnke, W. & Winfree, A. T. *Nature* **336**, 662–665 (1988).
13. Winfree, A. T. in *Chemical Waves and Patterns* (eds Kaplan, R. & Showalter, K.) Ch. 1 (Kluwer, Dordrecht, 1994).
14. Winfree, A. T. *When Time Breaks Down* (Princeton Univ. Press, 1987).
15. Panfilov, A. V. & Pertsov, A. M. *Proc. natn. Acad. Sci. USSR* **274**, 1500–1503 (1984).

16. Winfree, A. T. & Guilford, W. in *Biomathematics and Related Computational Problems* (ed. Riccardi, L. M.) 697–716 (Kluwer Academic, Dordrecht, 1988).
17. Winfree, A. T. & Strogatz, S. H. *Physica* **D13**, 221–233 (1984).
18. Keener, J. P. *Physica* **D31**, 269–276 (1988).
19. Winfree, A. T. & Strogatz, S. H. *Nature* **311**, 611–615 (1984).
20. Courtemanche, M., Skaggs, W. & Winfree, A. T. *Physica* **D41**, 173–182 (1990).
21. Henze, C., Lugosi, E. & Winfree, A. T. *Can. J. Phys.* **68**, 683–710 (1990).
22. Winfree, A. T. *Soc. ind. appl. Math. Rev.* **32**, 1–53 (1990).
23. Markus, M. & Hess, B. *Nature* **347**, 56–58 (1990).
24. Tyson, J. J. & Strogatz, S. H. *Int. J. Bifurc. Chaos* **1**, 723–744 (1991).
25. Gerhardt, M., Schuster, H. & Tyson, J. J. *Physica* **D50**, 189–206 (1991).
26. Henze, C. & Winfree, A. T. *Int. J. Bifurc. Chaos* **1**, 891–922 (1991).
27. Winfree, A. T. *Physica* **D49**, 125–140 (1991).
28. Winfree, A. T. in *1992 Lectures in Complex Systems* (eds Nadel, L. & Stein, D.) 207–298 (Santa Fe Inst. Stud. in the Sciences of Complexity, Addison-Wesley, Reading, Massachusetts, 1993).
29. Henze, C. thesis, Univ. Arizona (1993).
30. Nandapurkar, P. J. in *Simulation of Wave Processes in Excitable Media* (ed. Zykov, V. S.) (Manchester Univ. Press, 1988).
31. Winfree, A. T. *Chaos* **1**, 303–334 (1991).
32. Braune, M. & Engel, H. *Chem. Phys. Lett.* **204**, 257–264 (1993).
33. Nagy-Ungvári, Zs., Ungvári, J. & Müller, S. C. *Chaos* **3**, 15–19 (1993).
34. Jahnke, W. & Winfree, A. T. *Int. J. Bifurc. Chaos* **1**, 445–466 (1991).
35. Chen, P. S. et al. *Circulation Res.* **62**, 1191–1209 (1988).
36. Frazier, D. W. et al. *J. clin. Invest.* **83**, 1039–1052 (1989).
37. Winfree, A. T. in *Oscillations and Traveling Waves in Chemical Systems* (eds Field, R. J. & Burger, M.) 441–472 (Wiley, New York, 1985).
38. Winfree, A. T. *Prog. theor. Chem.* **4**, 1–51 (1978).

ACKNOWLEDGEMENTS. This work was performed during sabbatical leave from the University of Arizona in Tucson, with support from the US National Science Foundation and the Donors of the Petroleum Research Foundation of the American Chemical Society. Computations were carried out at the San Diego National Supercomputer Center.

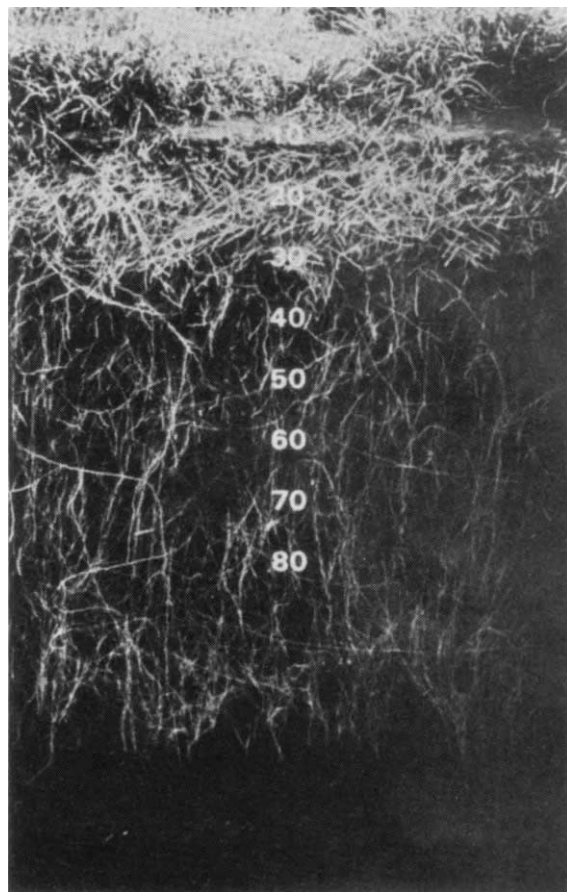


FIG. 1 *Andropogon gayanus* has the capacity to penetrate deep into oxisol soils of the savannas of tropical South America (from Spain and Couto¹⁷). The numbers indicate the soil depth in cm.

Carbon storage by introduced deep-rooted grasses in the South American savannas

M. J. Fisher, I. M. Rao, M. A. Ayarza, C. E. Lascano, J. I. Sanz, R. J. Thomas & R. R. Vera

Centro Internacional de Agricultura Tropical, Apartado Aéreo 6713, Cali, Colombia

ESTIMATES of the global carbon dioxide balance have identified a substantial 'missing sink' of 0.4–4.3 Gt per year¹. It has been suggested that much of this may reside in the terrestrial biosphere². Here we present an analysis of the carbon stored by pastures based on deep-rooted grasses which have been introduced in the South American savannas. Although the deep-rooted grasses were chosen principally for agricultural reasons³, we find that they also sequester significant amounts of organic carbon deep in the soil. If our study sites are representative of similar pastures throughout South America, this process could account for the sequestration of 100–507 Mt carbon per year—a substantial part of the 'missing sink'. Thus, although some land-use changes⁴ (such as burning tropical rainforests) contribute to the atmospheric CO₂ burden, we conclude that the introduced pastures studied here help to offset the effect of anthropogenic CO₂ emissions.

Savannas occupy some 250 million hectares (Mha) of South America, mainly in Brazil (200 Mha), Colombia (20 Mha) and Venezuela (12 Mha). They are used for extensive cattle ranching on the native forage, although in Brazil cropping with maize and soybeans (now 12 Mha) and introduced pastures (35 Mha) have become important during the past 30 years (ref. 5). The soils of the savannas are mainly oxisols and ultisols, characterized by low pH (4.0–4.8), aluminium saturation up to 90% and low levels of phosphorus and calcium. Rainfall is 1,500–3,000 mm with a unimodal seasonal distribution⁶.

The perennial grasses *Andropogon gayanus* and *Brachiaria humidicola* are of African origin^{7,8}; the former is tall-growing with a tussock habit, whereas the latter forms swards. The legumes *Arachis pintoii* and *Stylosanthes capitata* are from South America^{9,10}; *A. pintoii* is a vigorous, stoloniferous perennial and *S. capitata* is a free-seeding biannual. Grass–legume associations

of these forages produce cattle liveweight gains up to 500 kg ha⁻¹ yr⁻¹ (refs 11, 12), compared with 7–20 kg ha⁻¹ on well-managed savanna^{13,14}. They produce green forage for several months into the dry season, and regrow vigorously soon after the first rains of the wet season. Since 1980, all four have been released as cultivars (cv.) in one or more countries in South America¹⁵.

Although deep-rootedness is considered a major factor in adaptation to low-fertility soils, especially in *A. gayanus*^{16,17} (Fig. 1), the role of the roots in the dynamics of soil C has largely been ignored. We measured soil C, including fine roots, in an *A. gayanus* pasture, and in two *B. humidicola* pastures at two sites on the eastern plains (Llanos orientales) of Colombia, some 200 km apart (Table 1). In each case we obtained corresponding measurements from immediately adjacent native savanna with the same soil texture.

The pastures had differing histories (Table 1), but all had carried grass-based pastures for 3–9 years. The pasture at Matazol farm was not fertilized after the rice that was used as a pioneer crop to establish it, whereas at Carimagua fertilizer was applied at pasture establishment and each second year thereafter. They were all grazed by cattle at normal stocking rates for improved pastures in the region.

All the grass-based pastures made a striking contribution to soil C compared to the native savanna, especially when grown with a legume (Table 2). Data for other pastures based on another grass of African origin, *Brachiaria dictyoneura* cv. Llanero at Matazol farm, show lower, but still significant, sequestration of ~30 tons of carbon per hectare (t C ha⁻¹) in 3.5 years (data not presented). If these data are representative

TABLE 1 Location and characteristics of the two sites on the savannas of the eastern plains of Colombia

Site	Matazul farm	Carimagua research station
Location	Eastern plains (Llanos), Puerto Lopez, Colombia	Eastern plains (Llanos), 200 km ENE of Puerto Lopez, Colombia
Latitude, longitude	4° 9' N, 72° 39' W	4° 37' N, 71° 19' W
Altitude (m)	160	175
Mean annual rainfall (mm yr ⁻¹)	2,700	2,240
Soil	Oxisol	Oxisol
Texture	Clay loam	Clay loam
pH (1:1 water)	4.4	4.1
P (0–20 cm), Bray II (p.p.m.)	1.8	1.5
Pasture details	1989. Cropped from savanna with upland rice undersown with mixed <i>A.</i> <i>gayanus</i> cv. Carimagua 1 and <i>S. capitata</i> cv. Capica pasture. 1989–93. Rotationally grazed with cattle at 2 head per hectare.	1984. Sown to <i>B.</i> <i>humidicola</i> cv. Humidicola from savanna, with the legume <i>Desmodium</i> <i>ovalifolium</i> , which failed. 1987. Resown to <i>B.</i> <i>humidicola</i> cv. Humidicola alone or with <i>A. pintoi</i> cv. Mani Forrajero. 1988–93. Rotationally grazed with cattle at 3 head per hectare.
Date soil sampled	December 1992	April 1993

of the area sown to pastures of *A. gayanus* and *Brachiaria* species in South America, conservatively estimated at 35 Mha (ref. 5), then from 100 to as much as 507 Mt C is being sequestered each year.

The contribution of the legume in the 6 years since its establishment at Carimagua may be estimated by taking the difference between the grass alone and the association. This difference is 44.7 t C ha⁻¹, so that although legumes contribute only ~20% to root biomass¹⁸, the association with *A. pintoi* increased carbon sequestration by 7.8 t ha⁻¹ yr⁻¹ compared with the pure grass.

Compared with the savanna, the grass-based pastures sequester most of the C in the deeper part of the soil profile, well below the plough layer (normally 10–15 cm). This C should therefore be less prone to oxidation, and hence loss, during any

cropping phase that might be undertaken in integrated crop and pasture systems. Indeed, such systems should be able to accommodate rotations with annual crops and still contribute to C sequestration. Jones *et al.*¹⁹ drew attention to the role of fire in determining the balance between the vegetation of native savannas as either a net sink or source of carbon in the tropics. Introduced pastures are rarely burned, except by accident, in contrast to the native savannas, which are usually burned as often as each year.

There is little information about soil C below a depth of 15–20 cm in the tropics, although there are occasional measurements down to 40 cm. Long *et al.*²⁰ carried out careful studies on native tropical grasslands to document total primary productivity, including roots in the surface 15 cm. They showed that productivity of tropical grasslands, above ground and below ground to 15 cm depth, was up to five times higher than previously reported, mainly because losses due to senescence were ignored. On the basis of our data even they may have substantially underestimated root production. They did not present information on soil C.

Although the emphasis on deep-rooted grasses by the Centro Internacional de Agricultura Tropical was for reasons other than sequestration of C, we have shown that some introduced grasses in the tropical South American savannas do sequester C deep in the soil. There must be a physiological cost in growing such a massive root system, but we do not see much evidence of it. Both *A. gayanus* and *B. humidicola* grow at least as vigorously as other promising introduced grasses, and much more vigorously than the savanna species.

We suggest that the sequestration of C in South American savanna soils is of global importance, in that deep-rootedness could be exploited from the point of view of both the individual farmer and the community at large for their mutual benefit. Can the ability of tropical grasses to sequester C at depth, especially when grown in mixtures with legumes, be used in selection and breeding? What are the implications for the generation of technologies to recuperate degraded pastures in cleared areas of rainforests? The latter is an emotional issue, but there is considerable capacity for these soils to sequester C if they are managed correctly^{21,22}.

TABLE 2 Carbon in pastures compared to savanna

Site	Matazul farm			Carimagua research station				
	Pasture	<i>A. gayanus</i> /S. <i>capitata</i>		Pasture	<i>B. humidicola</i> alone		<i>B. humidicola</i> /A. <i>pintoi</i>	
Depth (cm)	Carbon in layer (t ha ⁻¹)	Carbon in layer (t ha ⁻¹)	Difference from savanna (t ha ⁻¹)	Carbon in layer (t ha ⁻¹)	Carbon in layer (t ha ⁻¹)	Difference from savanna (t ha ⁻¹)*	Carbon in layer (t ha ⁻¹)	Difference from savanna (t ha ⁻¹)*
0–20	64.0	71.1	7.1 ± 2.0†	70.3	76.0	5.7 ± 4.3‡	88.1	17.8 ± 4.2†
20–40	42.7	51.9	9.3 ± 2.8†	52.4	57.6	5.3 ± 3.2‡	71.2	18.6 ± 6.0†
40–100¶	79.8	114.2	34.3 ± 9.3§	74.3	89.2	14.9 ± 6.2	108.4	34.0 ± 10.0†
Total	186.5	237.2	50.7 ± 11.4§	197.0	222.8	25.9 ± 7.7†	267.7	70.4 ± 15.5§

Each sample comprised eight cores at Matazul and four at Carimagua, taken at random by soil auger to the depths indicated. In the grazed *A. gayanus*/S. *capitata* pastures at Matazul farm, samples were taken from each quarter of each of three 1-ha plots in a randomized complete block experiment covering 9 ha. At the same time four samples were taken from the native savanna immediately adjacent to the experiment. At Carimagua, the grass-based pastures were 0.5-ha plots in a randomized complete block grazing experiment. Five samples were taken from each of two replicates of the plots and three from the immediately adjacent savanna. Analysis methods were as follows. The samples were dried and milled to pass a 1-mm sieve. Fine roots were not removed before milling. Subsamples were digested in sulphuric acid/potassium dichromate heated (150 °C) for 30 min on a temperature-controlled hotplate. Carbon concentration in the digest was determined colorimetrically against calibrated standards made up of carbon-free soil to which measured amounts of glucose were added. Soil bulk density, used to convert the gravimetric figures of soil C to volumetric data, was determined by standard methods²⁵. The data for each depth of each pasture were treated as independent samples, and the standard error of each mean and the standard errors of the difference between the means were calculated. Differences between the means were tested for statistical significance with Student's *t*-test.

* Differences quoted are ±s.e., where s.e. is the standard error of difference between the means (*n* = 14 for Matazul farm, *n* = 12 for Carimagua).

Levels of significance indicated by symbols as follows:-

† *P* < 0.01. ‡ *P* > 0.05. § *P* < 0.001. || *P* < 0.05.

¶ At Carimagua research station the deeper layer was 40–80 cm.

The combination of a deep-rooted grass with a nitrogen-fixing legume can increase nutrient cycling, greatly improve animal production and markedly increase soil biological activity^{18,23}. These effects occur mainly at the soil surface while C storage takes place below the plough layer. Thus, far from being environmentally degrading, improved pastures can fulfill the restorative role in tropical systems that was recognized in pre-Roman times for Mediterranean systems²⁴, and may play a vital part in stabilizing the global carbon cycle and minimizing the greenhouse effect of atmospheric CO₂. □

Received 31 May; accepted 9 August 1994.

- Gifford, R. M. *Aust. J. Pl. Physiol.* **21**, 1–15 (1994).
- Watson, R. T., Rodhe, H., Oeschger, H. & Siegenthaler, U. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. & Ephraums, J. J.) 1–40 (Cambridge Univ. Press, Cambridge, 1990).
- Tropical Pastures Program Annual Report 1981* (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1982).
- Duxbury, J. M., Harper, L. A. & Mosier, A. R. in *Agricultural Ecosystem Effects on Trace Gases and Global Climate Change* 1–18 (Am. Soc. of Agronomy, Madison, 1993).
- Vera, R. R., Thomas, R., Sanint, L. & Sanz, J. I. *Anais Acad. bras. Cienc.* **64** (Supl. 1), 105–125 (1992).
- Cochrane, T. T., Sánchez, L. G., de Azevedo, L. G., Porras, J. A. & Garver, C. L. *Land in Tropical America*. Vol. 1 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1985).
- Ferguson, J. E., Vera, R. & Toledo, J. M. *Proc. XVI Int. Grassland Congr.* 1343–1344 (The French Grassland Society, Versailles, 1989).

- Thomas, D. & Grof, B. *Herbage Abstr.* **56**, 557–565 (1986).
- Valls, J. F. M. & Simpson, C. E. in *Biology and Agronomy of Forage Arachis* (eds Kerridge, P. C. & Hardy, B.) 1–18 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1994).
- Schultze-Kraft, R., Reid, R., Williams, R. J. & Coradin, L. in *The Biology and Agronomy of Stylosanthes* (eds Stace, H. M. & Edye, L.) 125–146 (Academic, Sydney, 1984).
- Lascano, C. L. & Thomas, D. in *Andropogon gayanus Kunth: A Grass for Tropical Acid Soils* (eds Toledo, J. M., Vera, R., Lascano, C. & Lenné, J. M.) 247–275 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1990).
- Lascano, C. E. in *Biology and Agronomy of Forage Arachis* (eds Kerridge, P. C. & Hardy, B.) 109–121 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1994).
- Fisher, M. J., Lascano, C. E., Vera, R. R. & Rippstein, G. in *Pastures for the Tropical Lowlands: CIAT's Contribution* 75–99 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1992).
- Paladines, O. & Leal, J. A. in *Pasture Production in Acid Soils of the Tropics* (eds Sánchez, P. A. & Tergas, L. E.) 311–325 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1979).
- Rao, I. M., Zeigler, R. S., Vera, R. & Sankarung, S. *BioScience* **43**, 454–465 (1993).
- Goedert, W. J., Ritchley, K. D. & Sanzonowicz, C. *Revta bras. Cienc. Solo* **9**, 89–91 (1985).
- Spain, J. M. & Couto, W. in *Andropogon gayanus Kunth: A Grass for Tropical Acid Soils* (eds Toledo, J. M., Vera, R., Lascano, C. & Lenné, J. M.) 223–246 (Centro Internacional de Agricultura Tropical, Cali, Colombia, 1990).
- Rao, I. M., Ayarza, M. A. & Thomas, R. J. *Pl. Soil* (in the press).
- Jones, M. B., Long, S. P. & Roberts, M. J. in *Primary Productivity of Grass Ecosystems* (eds Long, S. P., Jones, M. B. & Roberts, M. J.) 212–255 (Chapman & Hall, London, 1992).
- Long, S. P. *et al. Pl. Soil* **115**, 155–166 (1989).
- Lugo, A. E. & Brown, S. *Pl. Soil* **149**, 27–41 (1993).
- Veldkamp, E. *Soil Sci. Soc. Am. J.* **58**, 175–180 (1994).
- Thomas, R. J., Fisher, M. J., Ayarza, M. A. & Sanz, J. I. *Adv. Soil Sci.* (in the press).
- Hillel, D. *J. Out of the Earth* (Free Press, New York, 1991).
- Danielson, R. E. & Sutherland, P. L. in *Methods of Soil Analysis, Part 1: Physical and Mineralogical Methods* (ed. Klute, A.) 443–461 (American Society of Agronomy/Soil Science Society of America, Madison, 1986).

ACKNOWLEDGEMENTS. We thank P. M. Clausen for helpful comments on the manuscript and W. Scowcroft for encouragement.

Fluid trapping in mid-crustal reservoirs by H₂O–CO₂ mixtures

R. C. Bailey

Geology and Physics Departments, University of Toronto, Toronto, Canada M5S 1A7

A THIN reservoir of free aqueous or carbonic fluids at mid-crustal levels, perhaps in the form of water sills¹, has been suggested as the source of fault-transmitted metasomatizing fluids^{2–4}, as the cause of observed deep crustal electrical conductivity and anomalous seismic reflectivity⁵, and as a lubricated potential detachment zone for thin-skin tectonics⁶. A problem with this hypothesis is that estimated crustal permeabilities are too high to permit long-term retention of such fluids^{7,8}. Most mechanisms suggested for achieving the required permeability reduction rely on maintaining unusually low porosity^{2,6,9,10}. Because porosity creation, by tectonic deformation, fluid release or infiltration, is a ubiquitous process, permeability reductions achieved by these mechanisms are hard to sustain for long periods of time. A sealing (permeability reduction) mechanism that can operate in the presence of significant porosity is required. Here I propose that the required mechanism is capillary sealing by immiscible CO₂–H₂O mixtures derived from rising volatiles.

Several sources are possible for such rising volatiles. In or adjacent to tectonically active regions, prograde metamorphism at crustal depths can provide them. Another source of such volatiles is the upper mantle^{11,12}. Fyfe¹³ has noted that the entire contents of the oceans may be recycled by subduction in less than 10⁹ years. Most of this fluid is likely to be released for upward migration in, or adjacent to, tectonically active regions. Continuous fluid streaming through stable lower crust is improbable, however, because of the petrological evidence for a normally dry lower crust¹⁴. To be consistent with known lower crustal petrology fluid sources for sustained mid-crustal reservoirs must either be episodic, so that lower crust normally returns to a dry state, or involve horizontal migration at mid-crustal levels from active regions into stable regions, so that the lower crust is not involved. Such lateral migration is possible if a mid-crustal sealing mechanism operates as suggested in this Letter. To illustrate the role of the proposed sealing mechanism,

I consider the former scenario of a 'pulse' of metamorphic fluid supplied from the deep crust, but the arguments for sealing apply equally to the case of subhorizontally migrating water.

Driven by buoyancy forces, a pulse of supplied fluids will migrate rapidly upwards through ductile lower crust⁶, followed upward by prograding isograds, leaving the lower crust restored to its usual dry state¹⁵. On reaching the brittle–ductile transition, above which ductile porosity creation by pore inflation cannot occur, fluids may escape through fractures to the upper crust. But in a tectonically stable (or stabilizing) region, it seems unlikely that pre-existing fracture permeabilities will be important at depths just above the brittle–ductile transition; as their last tectonic experience, these rocks probably underwent the annealing associated with slow post-tectonic cooling at depth. Arrival of higher-pressure fluids from below would be expected to induce hydraulic fracturing. Such fracture planes would be subhorizontal in the normal stable continental stress regime of horizontal compression. The vertical extent of such a reservoir of hydraulically connected sills must be small; it cannot exceed the value over which the difference in hydrostatic and lithostatic gradients will produce a fluid overpressure sufficient to induce hydraulic fracturing. Secor and Pollard¹⁶ have shown theoretically that open fluid-filled vertical fractures in granite cannot extend vertically over more than a few hundred metres without initiating fracture extension. One might reasonably expect the same order of magnitude for maximum vertical reservoir extent to apply to horizontally extending hydraulic fractures. For mid-crustal depths, this number is a significant overestimate; the analysis in ref. 16 was for cold upper-crustal rocks. Significantly smaller strengths for rocks just above the brittle–ductile transition might be expected, and reservoir heights which are therefore significantly less than a few hundred metres.

The main difficulty with this hypothesis is that, even in the absence of vertical fracture permeability, the permeability of the sill roof rocks is likely to permit loss of water upwards in a geologically short time. Magnetotelluric evidence suggests an integrated thickness of water as sills (2–40 m)⁶. This would be lost upwards in a geologically short time if rock permeabilities exceed ~10^{–23} m². This value is at the lower limit of measured permeabilities of intact rock specimens, and very much lower than estimated bulk permeability of crystalline rocks⁷. As a solution to this difficulty, I propose that the formation of immiscible

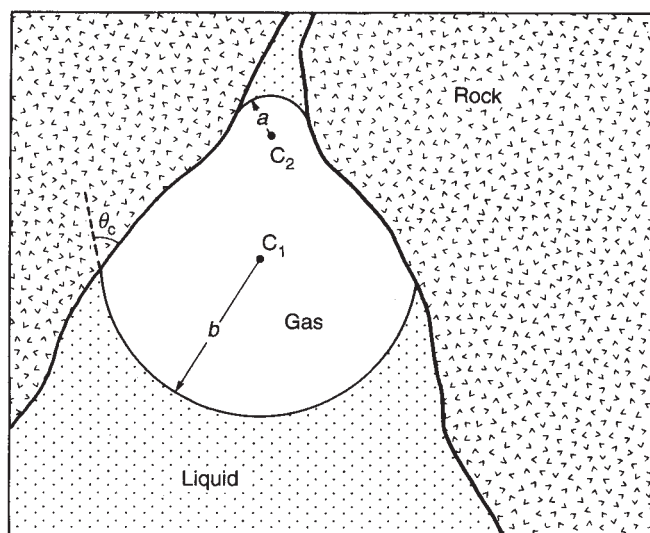


FIG. 1 Constricting pore in which a trapped gas bubble prevents further fluid flow. Pore radii at bubble bottom (b) and top (a) are shown, along with the contact angle θ_c . C_1 and C_2 are nominal centres of curvature of the two end surfaces.

CO₂-H₂O mixtures in the roof rocks can seal the porosity against upward migration. Although the importance of H₂O-CO₂ immiscibility to crustal fluid flow is well known^{11,17}, the depth to which such immiscibility persists is less well known. Experimental results^{18,19} and theoretical²⁰ calculations have suggested that the systems H₂O-NaCl-CO₂ and H₂O-CaCl₂-CO₂ are immiscible in at least the upper third of the crust to temperatures of 700 °C. Recent experiments²¹ show that H₂O-NaCl-CO₂ immiscibility in granulites exists at 7 kbar and 930 °C for moderate NaCl and CO₂ contents, and concludes that the required salinities are reasonable in terms of salinities reported in the literature for high-grade metamorphic rocks. Any reasonable interpolation of these data suggests that CO₂ and deep-crustal brines (which are often dominated by Ca rather than Na²²) are immiscible for at least some CO₂-rich compositions to depths below the brittle-ductile transition, and quite possibly to the base of the crust in some cases.

For such two-phase fluids, permeability reduction mechanisms associated with capillary effects are well known in hydrology and petroleum reservoir engineering²³. Consider a bubble of the non-wetting phase (CO₂) forced into a pore constriction (throat) by a pressure gradient in the fluid (see Fig. 1). The pressure drop Δp between the liquid phase at the bottom and top ends of the bubble is given by Laplace's equation of capillarity

$$\Delta p = 2\sigma \left(\frac{1}{b} - \frac{1}{a} \right) \cos \theta_c \quad (1)$$

in terms of the interfluid surface tension σ , the wall-interface contact angle θ_c and the pore body and throat radii b and a respectively. $\cos \theta_c$ can be here approximated as of order unity. If the average pore size (throat-to-throat distance) is of order d , and all pore throats are blocked by bubbles, then a mean pressure gradient in bulk rock of

$$\nabla p = \frac{2\sigma(1-f)}{bd} \quad (2)$$

can be maintained, where f is the pore body-to-throat ratio b/a . This is the order of magnitude of the threshold pressure gradient which must be exceeded to produce fluid flow in the rock.

What upward fluid pressure gradients are available? In the deeper ductile crust, fluid pressures must be lithostatic. In shallow crust, they are hydrostatic. In the transition zone, the fluid

pressure gradient must necessarily exceed the lithostatic gradient, perhaps significantly. Thus in order to trap fluids by the mechanism suggested here, the threshold pressure gradient given by equation (2) must be able to significantly exceed the lithostatic gradient.

The geometry shown in Fig. 1 is oversimplified. Real pore surfaces are microscopically rough, and the wetting phase cannot be entirely displaced using any finite pressure from the bottoms of micro-valleys which are nominally covered by the non-wetting phase. Dullien *et al.*²⁴ demonstrate experimentally that in high-porosity materials such as Berea sandstone, these wetted micro-valleys form a connected network around the non-wetting phase bubbles. But these experiments were done with very strongly wetting fluids, and Katz and Trugman²⁵ have argued that such connection may not occur for the less strongly wetting fluids encountered in the Earth. Strong wetting by water in the middle to deep crust seems unlikely. Although room-temperature measurements of contact angles show that distilled water wets clean quartz surfaces perfectly, this is not true for oligoclase²⁶. At high temperature (1,000 °C), measurements of dihedral angles of saline water against quartz, olivine, oligoclase, orthopyroxene and synthetic diopside give values all exceeding 40° (ref. 27).

Application of equation (2) requires an estimate of surface tension for deep crustal conditions. The surface tension of pure water at STP is 0.073 N m⁻¹. The addition of simple ionic salts such as NaCl raises the surface tension slightly, to 0.078 N m⁻¹ in the case of 20 wt% NaCl at STP (ref. 28). As the temperature is raised in a two-phase fluid, the surface tension falls to zero at the critical temperature T_c ; the descent is fairly robustly described by a critical exponent μ of ~ 1.25 (ref. 29), so that

$$\sigma(T) \approx \sigma_0 \left(1 - \frac{T}{T_c} \right)^\mu \quad (3)$$

where σ_0 is the low temperature limit of σ . As noted above, critical temperatures for plausible crustal fluids may exceed 900 °C; it is therefore reasonable to postulate surface tensions for a brine-CO₂ interface of order 0.05 N m⁻¹ down to mid-crustal depths at least. This surface tension estimate can be used to estimate threshold pressure gradients for fluid flow in various rock types.

For grain-scale porosity, results for Westerly granite under pressure³⁰ suggest values for d and b of 500 and 0.2 μ m, respectively; 2 would be an extremely conservative estimate of f . The resultant threshold pressure gradient is $\sim 10^9$ Pa m⁻¹, or 30,000 times lithostatic. For the more porous Sherman granite³⁰, b is ~ 12 μ m; this still gives a threshold more than 1,000 times lithostatic. Even for a very porous upper-crustal rock such as a sandstone, for which b might be as high as half of d (250 μ m) the threshold gradient is still 30 times lithostatic. The minimum required sealing zone thickness $h_{\min}$ can be calculated by equating the above threshold gradients to the gradient associated with a transition from lithostatic to hydrostatic pressure across the sealing thickness. For these three rocks, at a depth of 10 km, $h_{\min}$ is approximately 25 cm, 7 m and 250 m, respectively. There is clearly no problem in sealing any reasonable middle- or deep-crustal rock against upward fluid migration if the two-phase mixture is sufficiently dispersed to clog almost all available pores.

Achieving this uniform dispersal is not possible by simply sweeping the fluid into the rock; the non-wetting phase will only occupy those pores into which the available pressure gradient will sweep it. Connected pathways through smaller pores will exist to leak the wetting phase. Therefore a suitable dispersed mixture must be generated *in situ*, either by co-generating the two fluid phases locally (for example, by prograde metamorphism), or by unmixing the two phases as they rise upwards to lower pressures and temperatures. Although the former mechanism must also occur, only the latter mechanism can account for widespread trapping.

Consider therefore the evolution of a rising CO₂-brine fluid on the idealized pseudobinary phase diagram shown in Fig. 2, which shows the single-phase/two-phase boundary as a function of temperature (and thus effectively depth). For rising H₂O-rich fluids (case A in Fig. 2) the very small non-wetting CO₂ bubbles exsolved in the fluid will simply rise through the porous network until trapping in the manner described above, either against a pore throat or against a previously trapped bubble. The eventual lengthening of bubbles thus produced to dimensions much longer than a grain size will reduce the threshold pressure gradient to below the actual driving pressure gradient, and the bubbles will move again. This coalescence of bubbles will prevent the dispersal required by the trapping mechanism described above.

On the other hand, if CO₂-rich fluids are rising (case B in Fig. 2), water will be exsolved as the minority constituent; because it is the wetting phase, it will adhere to mineral surfaces in thin lenses. When water accumulation increases enough at pore throats, the phenomenon of choke-off³¹ will block pore throats with water; rising CO₂ will be compartmentalized. Continued percolation of fluid through those parts of the network still open will in turn seal those parts. Thus, rising water-saturated CO₂ can eventually generate an immobile two-phase mixture.

Such sealing cannot be permanent below the brittle-ductile transition. Over long timescales at greater depth, ductile deformation of the rock matrix around the pore space will occur. Because the gas pressure is higher than the liquid pressure (by $\sim 10^{-3}$ bar for a bubble of 10^{-4} m radius), the rock will retreat from rock-gas interfaces and advance on rock-liquid interfaces, eventually leading to the isolation of the gas bubble inside the liquid away from rock surfaces, permitting easy liquid leakage around at least some parts of the bubble. For calculating deep-crustal water residence times, CO₂ can be regarded as part of the deformable matrix through which the water must flow; these times are geologically short⁶. Thus although CO₂ can remain trapped at greater depths, water cannot, consistent with current petrological views of the lower crust¹⁴. As sealing cannot occur below the brittle-ductile transition, it would probably be most pronounced at those depths where it first becomes possible for rising fluids, namely just above the brittle-ductile transition. At this depth, the fluid will pond from below against the sealing level.

Can such a process generate horizontally extended reservoirs, given the horizontal variability of rock properties and fluid

inputs? Probably yes, because the proposed trapping mechanism is self-healing. If by horizontal spreading, ponding fluid below the sealing level reaches an unsealed region, and leaks upward at that point, the water-exsolution process described above can generate a new seal. The final state of the reservoir zone might well resemble an inverted version of puddles on uneven ground after a rainfall. In exhumed rocks, these would presumably show as originally subhorizontal fractures (for faults and shear zones, if exhumation was associated with tectonic stress) with associated local retrogression. □

Received 18 March; accepted 27 July 1994.

1. Fyfe, W. S., Price, N. J. & Thomson, A. B. *Fluids in the Earth's crust* (Developments in Geochemistry Vol. 1, Elsevier, New York, 1978).
2. Etheridge, M. A., Wall, V. J., Cox, S. F. & Vernon, R. H. *J. geophys. Res.* **89**, 4344-4358 (1984).
3. Sibson, R. H. *Geology* **15**, 701-704 (1987).
4. Rice, J. R. in *Fault Mechanics and Transport Properties of Rocks* (eds. Evans, B. & Wong, T.-F. 475-503 (Academic, New York, 1992).
5. Jones, A. G. *Geophys. J.* **89**, 7-18 (1987).
6. Bailey, R. C., *Geophys. Res. Lett.* **17**, 1129-1132 (1990).
7. Brace, W. F. *Int. J. Rock Mech. Miner. Sci. Geomech. Abstr.* **17**, 241-251 (1980).
8. Hyndman, R. D. & Shearer, P. M. *Geophys. J. Int.* **98**, 343-365 (1989).
9. Walder, J. & Nur, A. J. *Geophys. Res.* **89**, 11539-11548 (1984).
10. Marquis, G. & Hyndman, R. D. *Geophys. J. Int.* **110**, 91-105 (1992).
11. Thomson, A. B. *Nature* **358**, 295-302 (1992).
12. Sibson, R. H., *J. struct. Geol.* **11**, 873-877 (1989).
13. Fyfe, W. S. in *Saline Water and Gases in Crystalline Rocks* (eds Fritz, P. & Frape, S. K.) 1-3 (Spec. Pap. No. 33, Geol. Ass. Canada, Ottawa, 1987).
14. Yardley, B. W. D. *Nature* **323**, 111 (1986).
15. Thompson, A. B. & Connolly, J. A. D. *Tectonophysics* **182**, 47-55 (1990).
16. Secor, D. T. & Pollard, D. D. *Geophys. Res. Lett.* **2**, 510-513 (1975).
17. Yardley, B. W. D. & Bottrell, S. H. *Geophys. J.* **16**, 199-202 (1988).
18. Frantz, J. D., Popp, R. K. & Hoering, T. C. *Chem. Geol.* **98**, 237-255 (1992).
19. Zhang, Y. G. & Frantz, J. D. *Chem. Geol.* **74**, 289-308 (1989).
20. Bowers, T. S. & Helgeson, H. C. *Geochim. cosmochim. Acta* **47**, 1247-1275 (1983).
21. Johnson, E. L. *Geology* **19**, 925-928 (1991).
22. Frape, S. K. & Fritz, P. in *Saline Water and Gases in Crystalline Rocks* (eds Fritz, P. & Frape, S. K.) 19-37 (Spec. Pap. No. 33, Geol. Ass. Canada, Ottawa, 1987).
23. Stegemeier, G. L. in *Improved Oil Recovery by Surfactant and Polymer Flooding* (eds Shah, D. O. & Schechter, R. S.) 55-91 (Academic, New York, 1977).
24. Dullien, F. A. L., Lai, F. S. Y. & MacDonald, I. F. J. *Colloid Interface Sci.* **109**, 201-218 (1986).
25. Katz, A. J. & Trugman, S. A. *J. Colloid Interface Sci.* **123**, 8-13 (1988).
26. Nassau, K. & Schonhorn, H. *Gems Gemol.* **15**, 354-360 (1978).
27. Watson, E. B. & Brennan, J. M. *Earth planet. Sci. Lett.* **85**, 497-515 (1987).
28. Adamson, A. W. *Physical Chemistry of Surfaces* 70 (Interscience, New York, 1960).
29. Rowlinson, J. S. & Widom, B. *Molecular Theory of Capillarity* (Clarendon, Oxford, 1982).
30. Brace, W. F., Walsh, J. B. & Frangos, W. T. *J. geophys. Res.* **73**, 2225-2236 (1968).
31. Mohanty, K. K., Davis, H. T. & Scriven L. E. in *Proc. Soc. Petrol. Engrs 55th A. Conf.* (Publ. No. 9406, Soc. Petrol. Engrs, Dallas, 1980).

Parental choice selects for ornamental plumage in American coot chicks

Bruce E. Lyon*†, John M. Eadie* & Linda D. Hamilton†

* Division of Life Sciences, University of Toronto, Scarborough Campus, Scarborough, Ontario M1C 1A4, Canada

† Faculty of Medicine, University of Calgary, 3330 Hospital Drive NW, Calgary, Alberta T2N 1N4, Canada

ORNAMENTAL traits such as colourful bird plumage were the prime motivation for Darwin's theory of sexual selection¹. Other evolutionary mechanisms could also select for ornamental traits²⁻⁶, but such mechanisms have received far less attention, and empirical evidence for their existence is weak. Here we show that parental choice selects for the bizarre ornamental plumes of newly hatched American coot (*Fulica americana*) chicks. Experimental manipulations of chick plumage revealed that parent coots feed ornamented chicks preferentially over non-ornamented chicks, resulting in higher growth rates and greater survival for ornamented chicks.

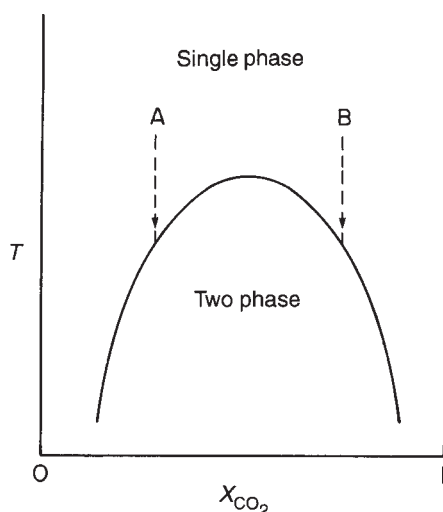


FIG. 2 Temperature defining the miscibility limit as a function of composition X_{CO_2} (mole fraction of CO₂) for an idealized pseudobinary phase diagram of the saline H₂O-CO₂ system. A and B are evolutionary paths for two possible fluid compositions described in the text.

† Present address: Kananaskis Field Stations, The University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada.

Moreover, we show that parental preference is relative, rather than absolute, an important element in the evolution of exaggerated traits^{7,8}. These observations provide the first empirical evidence that parental choice can select for ornamental traits in offspring.

American coots are monogamous rails (Rallidae) that breed throughout marshes in western North America. Both sexes share in territory defence, incubation, and feeding the precocial offspring for at least a month after hatching. American coot chicks are a striking exception to the general pattern of cryptic natal plumages in precocial chicks^{9,10}. Conspicuous orange, waxy-tipped filaments cover the front half of the body, brilliant red papillae surround the eyes and base of the bright red bill, and the top of the head is naked and brightly coloured (Fig. 1). The structure of the colourful plumes indicates that they are unlikely to play a role in thermoregulation (Fig. 1), and thus probably function strictly as signals.

The presence of conspicuous ornaments like colourful plumes and skin patches in newly hatched chicks is particularly astonishing because flightless chicks should be especially vulnerable to predators. Coot chicks frequently hide their heads when adults give alarm calls (personal observation), indicating that this colourful plumage does have associated costs. Although sexual selection would normally be invoked to explain the occurrence of potentially costly ornamental traits like conspicuous colouration^{8,11,12}, that mechanism cannot apply here because the colourful plumage is lost by the age of three weeks. However, a variety of social interactions other than sexual selection could potentially favour the evolution of exaggerated traits²⁻⁶. For example, West-Eberhard⁵ noted that parents are often in a position to exercise favouritism among their offspring through the allocation of critical resources like food. If parents preferentially feed offspring that bear the most extreme expression of a particular trait and if such feeding enhances offspring fitness, parental choice could result in the evolutionary elaboration of the offspring trait.

This idea, previously untested in any animal, may apply to the plumage of coot chicks. Parental feedings are essential for chick survival and between one third and one half of all chicks that hatch subsequently perish, mainly as a result of starvation¹³. Eggs hatch asynchronously and later hatched chicks suffer significantly higher mortality than early hatched chicks¹³. Aggres-

sion among chicks does not occur and parents exercise total control over the delivery of food among offspring (B.E.L., unpublished data). Beginning the week following hatching, brood division occurs whereby each parent specializes in feeding a subset of the brood. The youngest chicks attended by each parent are fed the most (B.E.L., unpublished). Thus, parents control the allocation of resources critical for chick survival and are in a position to exert choice. Chicks do not vocalize when begging but rather approach parents and display their plumage and, less frequently, the naked skin on the top of their head¹⁰. These displays suggest that parents assess chick plumage during feeding interactions and that parental choice may underlie the conspicuous plumage.

To test this hypothesis, we experimentally manipulated the plumage of chicks in a population of American coots in central British Columbia, Canada, during the breeding season of 1992. Our study population comprised 90 pairs of coots on three permanent wetlands (5–15 ha). A pilot study indicated that it was not feasible to dye chicks—the dye made chicks ill and removed waterproofing oils from their plumage. We therefore altered the appearance of chicks by trimming the orange tips off their body feathers. In each experimental brood, we trimmed half of the chicks in the brood, hereafter termed 'black' chicks, and left the remaining half of the brood with their orange plumes intact ('orange' chicks). On black chicks, the orange plumage was restricted to the throat area, a plumage coloration that closely resembles the natal plumage of most Rallids, including several congeners^{10,14}. Thus, our experimental manipulations produced an ancestral appearance rather than a novel one¹⁵.

With the exception of trimming, all chicks were handled in a similar manner and were held in captivity for about 30 min. Trimmed chicks did not appear appreciably smaller as a result of the removal of the orange feather tips. To ensure further that the removal of feathers *per se* did not affect the viability of chicks and thus confound our experiment, we used a double set of control broods. In black control broods, we trimmed the orange feather tips from all chicks in the brood, and in orange control broods all chicks were left with their natural plumage intact. If the orange plumes are strictly ornamental, chicks in these two control groups should not differ in any measure of parental care or fitness. We found no difference between the two control groups in any of the measures (Fig. 2a–c; Mann-Whitney *U* tests in all cases); feeding rates, $U' = 85$, $P = 0.24$; growth rates, $U' = 59$, $P = 0.78$, or chick survival, $U' = 91.5$, $P = 0.94$. The almost identical feeding rates and viabilities of chicks in the two control groups confirm the validity of our experiment and demonstrate that the colourful plumes are strictly ornamental.

Within experimental broods, orange chicks were fed at a higher rate than black chicks (Fig. 2d; Wilcoxon test, $z = 3.84$, $P < 0.001$), demonstrating that parents prefer ornamented chicks. As a result of their enhanced feeding rate, orange chicks grew more rapidly (Fig. 2e; Wilcoxon test, $z = 2.83$, $P < 0.005$) and enjoyed higher survival rates than black chicks in the same brood (Fig. 2f; Wilcoxon test, $z = 2.77$, $P < 0.01$). Parental preference clearly translates into increased fitness for ornamented offspring.

The survival benefits from increased parental care did not accrue to all orange chicks in experimental broods, but depended strongly on a chick's position in the hatching order (Fig. 3). Late-hatched orange chicks had dramatically higher survival than late-hatched black chicks, but there was little difference in survival between early-hatched black and orange chicks (Fig. 3). This result is predicted, given the patterns of parental care in American coots. After brood division occurs, each parent has one or two 'favourite' chicks that receive most of the feeding. These are primarily the late-hatched chicks (B.E.L., unpublished data). Accordingly, any influence of ornamentation on parental feeding should be most strongly evidenced in the late-hatched chicks, as we found. Ornamental plumes thus appear to increase

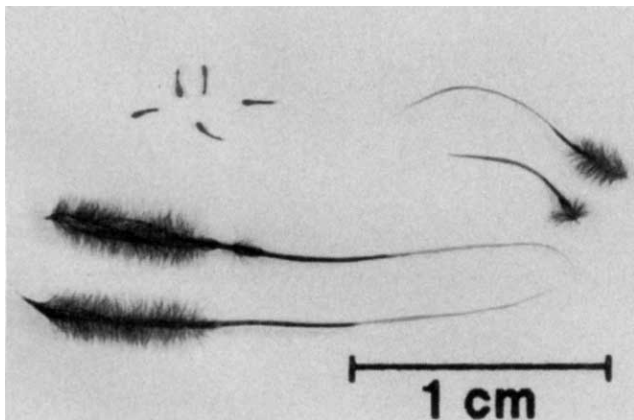


FIG. 1 Structure of the ornamental plumes from three regions of the body. The conspicuous coloration occurs on the modified distal ends of the throat feathers (upper right) and back feathers (bottom) while the black down typical of Rallid chicks is restricted to the basal portion. The reddish papillae-like feathers that surround the eyes and base of the bill (upper left) are highly modified and lack down completely. The structure of the colourful portions of these ornamental feathers implies a modification from a thermoregulatory function to a signalling function.

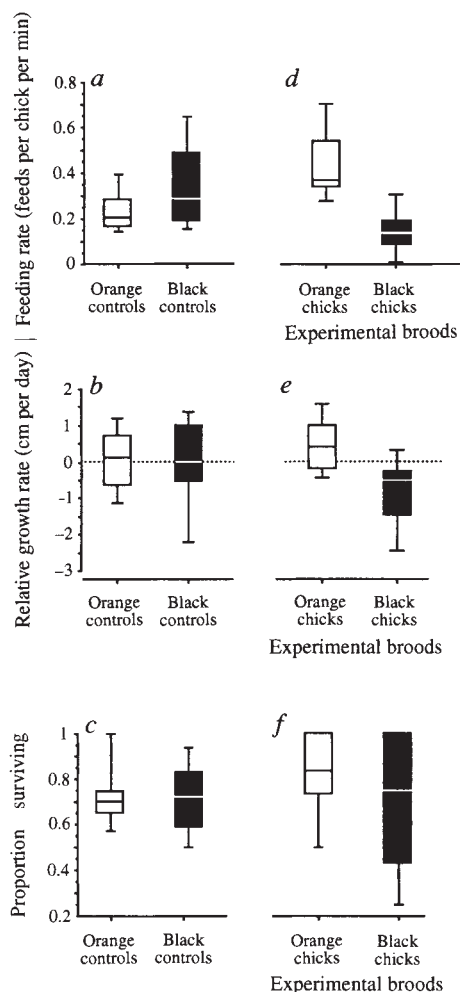


FIG. 2 Feeding rates, growth rates and survival of chicks in control (a–c) and experimental (d–f) broods. Values shown are medians, interquartile ranges, and 10–90 percentiles. To avoid pseudoreplication, a single mean value was calculated over all chicks in each control brood and each mean was then used as an independent datum in two sample comparisons. For experimental broods, mean values were calculated separately over all orange chicks and all black chicks, respectively, within each brood. Matched-paired comparisons were then used to compare these values for each treatment within broods. Plumage on chicks were trimmed within a day of hatching. To control for the confounding influence of hatching order²¹, the first chick to hatch in each experimental brood was randomly assigned to a treatment and thereafter treatments were alternated with hatching order. Chicks in each brood were individually colour-marked and broods were observed from floating blinds to estimate feeding rates, growth and survival²¹. All observation periods for each brood were pooled to calculate one overall mean feeding rate (average of 1.4 h observation per brood at 12 orange control broods, 11 black control broods, 21 experimental broods). To estimate growth rate, swimming chicks were photographed at known distances and body length at waterline was estimated from projected slides²². This measure of body size is strongly correlated with body mass²² ($r=0.97$, $n=43$). Residual values from a regression of body length on chick age were used as an index of relative growth rate, and mean brood values were calculated for the 11 orange control, 10 black control and 13 experimental broods. Survival comparisons are based on 15 orange control, 12 black control and 25 experimental broods. Sample sizes differ among comparisons because some broods could not be monitored for feeding and/or growth rates. Sixteen broods (9 experimental, 4 orange controls, 3 black controls) were excluded from the analysis because they adopted chicks or donated to chicks other broods, but the results are unchanged if these broods are included.

the likelihood that a chick will be chosen as a favourite chick once brood division takes place.

An important issue in the evolution of exaggerated traits through a mechanism such as female choice or parental choice is whether the preference is relative or absolute^{7,8}. Our experiment confirms that parental preference in coots is relative. Non-ornamented chicks were only disadvantaged when ornamented chicks were also present in the brood (experimental broods). In black control broods, for which there was no variation from which parents could choose, non-ornamented chicks ate, grew and survived as well as the ornamented chicks in orange control broods. This finding is important because it allows us to reject the hypothesis that non-ornamented chicks fare poorly in experimental broods simply because they are not recognized as coot chicks by their parents.

Given that parental preference is relative, an alternative explanation for the observed pattern is possible—perhaps trimming feathers somehow influenced chick begging behaviour so that parents responded to behavioural differences rather than plumage differences. Begging differences between trimmed and non-trimmed chicks might result in higher feeding rates for non-trimmed chicks in experimental broods, but no difference would be expected between the two groups of controls owing to a lack of variation within these broods. However, we could detect no difference in the behaviour of trimmed and non-trimmed chicks in observations conducted within a day or two of the trimming. Moreover, we cannot envisage a mechanism by which the removal of the thin, highly modified feather tips would affect chick begging behaviour sufficiently to produce the large differences observed in parental feeding rates. A second possibility is that trimming made black chicks appear smaller than orange chicks and parents responded to the apparent size of chicks rather than plumage colour. This hypothesis is rejected by the observation that parents preferentially feed the smallest chicks—black chicks should therefore have been fed more, not less, if this mechanism applied. We conclude that the trait under selection by parental choice is the ornamental plumage of coot chicks.

Our study now raises the question of why parents prefer ornamented chicks in the first place. We suggest three possibilities. First, ornamental plumes may provide an honest signal of a

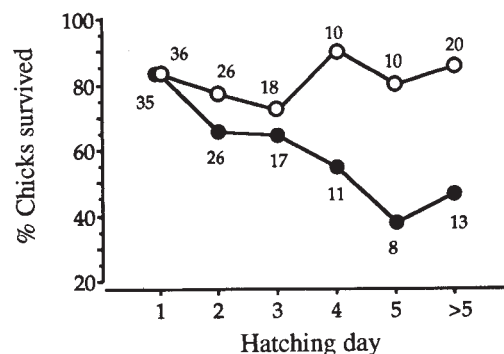


FIG. 3 The relationship between hatching day and survival of orange chicks (open circles) and black chicks (filled circles) within experimental broods. Hatching in American coots is asynchronous; day 1 represents the day on which the first chicks in a brood hatched. Sample sizes (number of chicks) are shown beside each point. Analysis of log likelihood revealed a significant interaction effect between hatching order and plumage colour on the survival of chicks (Wald $\chi^2=4.92$, d.f. = 1, $P<0.03$). Post-hoc comparisons indicate no effect of plumage colour on the survival of early chicks (days 1–3 pooled, G-test with Williams' correction, $G_w=0.69$, d.f. = 1, $P>0.50$ after Bonferroni correction), but a significant effect of plumage colour on the survival of late chicks (days 4 + pooled; $G_w=11.85$, d.f. = 1, $P<0.003$ after Bonferroni correction).

chick's genetic or phenotypic quality^{16,17} which allows parents to invest selectively in high-quality offspring¹⁸. For such a signal to be reliable, it must be costly^{16,17,19}; we cannot yet assess whether this might apply to the bright plumage of coot chicks. Second, the orange plumage may provide an honest signal of age, allowing parents to allocate food optimally among offspring. Finally, the preferences for the orange plumage may involve sensory exploitation of a parental colour preference favoured for other reasons²⁰. For example, the bare skin patch on the chick's head varies in colour over time and may signal off-

spring need¹⁹. At its brightest intensity, the skin is similar in colour to the ornamental plumes (B.E.L. *et al.*, unpublished data). Thus, the initial parental preference for orange colouration may have evolved as a response to a condition-dependent signal of the naked skin, with the elaboration of bright plumage following as a secondary consequence of this sensory bias. We are currently investigating these possibilities, although, as with studies of sexual selection, demonstrating that a preference exists may prove easier than understanding why the choice mechanism itself has been favoured by selection²⁰. □

Received 18 April; accepted 13 July 1994.

1. Darwin, C. *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).
2. Dobzhansky, T. *Am. Nat.* **74**, 312–321 (1940).
3. Baker, R. R. & Parker, G. A. *Trans. R. Soc. Lond. B* **287**, 63–130 (1979).
4. West-Eberhard, M. J. *Proc. Am. Phil. Soc.* **132**, 222–243 (1979).
5. West-Eberhard, M. J. *Q. Rev. Biol.* **58**, 155–183 (1983).
6. Zahavi, A. *Anim. Behav.* **42**, 501–503 (1991).
7. Arnold, S. in *Mate Choice* (ed. Bateson, P.) 67–107 (Cambridge Univ. Press, Cambridge, 1983).
8. Zuk, M., Johnson, K., Thornhill, R. & Ligon, J. D. *Evolution* **44**, 477–485 (1990).
9. Harrison, C. *A Field Guide to the Nests, Eggs and Nestlings of North American Birds* (Collins, Toronto, 1984).
10. Boyd, H. J. & Alley, R. *Ibis* **90**, 582–593 (1948).
11. Andersson, M. *Nature* **299**, 818–820 (1982).
12. Kodric-Brown, A. *Behav. Ecol. Sociobiol.* **17**, 199–205 (1985).

13. Lyon, B. *Anim. Behav.* **46**, 911–928 (1993).
14. Ripley, D. *Rails of the World* (Godine, Boston, 1977).
15. Ryan, M. J. & Keddy-Hector, A. *Am. Nat.* **139**, S4–S35 (1992).
16. Zahavi, A. in *International Symposium of Biological Evolution* (ed. Delfino, V. P.) (Adriatica Editrice, Bari, 1987).
17. Grafen, A. J. *Theor. Biol.* **144**, 517–546 (1990).
18. Stephenson, A. G. & Winsor, J. A. *Evolution* **40**, 453–458 (1986).
19. Godfray, H. C. J. *Nature* **352**, 328–330 (1991).
20. Kirkpatrick, M. & Ryan, M. J. *Nature* **350**, 33–38 (1991).
21. Lyon, B. thesis, Princeton Univ. (1992).
22. Lyon, B. *Condor* (in the press).

ACKNOWLEDGEMENTS. We thank R. Boonstra, M. Gross, G. Hill, W. Hochachka, K. Holder, D. Lank, R. Montgomerie and D. Queller for comments on the manuscript. B. Boyle and R. Cartar for discussion, and the Humphreys, whose cooperation made this study possible. This research was funded by an NSERC operating grant to J.M.E.

Electrostatic tuning of Mg^{2+} affinity in an inward-rectifier K^+ channel

Zhe Lu & Roderick MacKinnon

Department of Neurobiology, Harvard Medical School,
220 Longwood Avenue, Boston, Massachusetts 02115, USA

INWARD-RECTIFIER potassium channels conduct K^+ across the cell membrane more efficiently in the inward than outward direction. This unusual conduction property is directly related to the biological action of these channels^{1–6}. One basis for inward rectification is voltage-dependent blockade by intracellular Mg^{2+} (refs 1, 7–9): strong inward-rectifier channels are so sensitive to intracellular Mg^{2+} that no outward K^+ current is measurable under physiological conditions; weak inward rectifiers are less sensitive and allow some K^+ to flow outwards. Background $K1$ channels and acetylcholine-regulated K^+ channels from the heart are examples of strong inward rectifiers and ATP-sensitive K^+ channels are weak rectifiers^{1,7–10}. Here we show that mutations at one position in the second transmembrane segment can alter the Mg^{2+} affinity and convert a weakly rectifying channel (ROMK1) into a strong rectifier. The amino acid at this position exposes its side chain to the aqueous pore and affects Mg^{2+} blockade as well as K^+ conduction through an electrostatic mechanism.

Figure 1 illustrates two general features of Mg^{2+} blockade in an inward rectifier K^+ channel, ROMK1. First, blockade due to intracellular Mg^{2+} is highly voltage-dependent. The voltage dependence accounts for the curvature (inward rectification) in the current-voltage ($I-V$) graphs (Fig. 1a). Second, as the external K^+ concentration ($[K^+]_o$) is raised, Mg^{2+} affinity is decreased: at higher external $[K^+]_o$ stronger depolarization is required to achieve the same degree of blockade (Fig. 1b, c). Taken together, the voltage dependence and effect of extracellular $[K^+]_o$ argue that Mg^{2+} occludes the ion-conduction pore^{7–9}.

The membrane folding topology of inward-rectifier K^+ channels is unknown. The homology between voltage-activated and inward-rectifier K^+ channels suggests that inward rectifiers contain a unit consisting of a P-region flanked by two transmembrane segments^{11,13}. Based on mutagenesis studies in voltage-

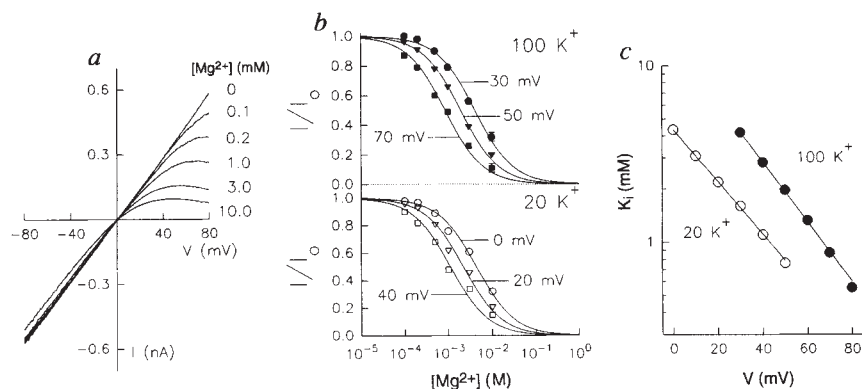
activated K^+ channels, part or all of the unit could be involved in ion conduction^{14–18}. Figure 2a shows a sequence alignment of the second membrane-spanning segment of three inward-rectifier K^+ channels^{11,13,19,20}. IRK1 and GIRK1 are strong rectifiers but ROMK1 is a weak rectifier^{11–13,19–21}. Figure 2 shows that ROMK1 becomes a strong rectifier when aspartate (D), the residue found in IRK1 and GIRK1, is substituted for asparagine (N) at position 171. In cell-attached patch recordings, the wild-type ROMK1 channel shows weak inward rectification (Fig. 2c). By contrast, the mutant channel with an aspartate substitution at position 171 (N171D) shows no outward current (Fig. 2e). The dramatic effect on the $I-V$ relationship can be attributed to the carboxyl side chain of aspartate because substituting glutamate (Fig. 2f), but not threonine (Fig. 2d), has an effect nearly identical to that of aspartate. In these cell-attached recordings, we assume that the affinity for intracellular Mg^{2+} has been significantly increased by placing an acidic residue at position 171.

Sixteen residues were substituted at position 171 and only those with hydrophilic side chains produced currents in oocytes (Fig. 2b). All were highly selective for K^+ over Na^+ (data not shown). The Mg^{2+} sensitivity of mutant channels in excised, inside-out patches is shown in Fig. 3. Both macroscopic and single-channel records demonstrate a 20-fold increase in Mg^{2+} affinity when Asp is present at position 171 (Figs 1a and 3). (In single-channel records Mg^{2+} reduces the amplitude because discrete blocking events are filtered by the relatively slow response time of our recording system (Fig. 3g).) Mg^{2+} affinity is altered by the same factor at all voltages, implying that the N171D mutation affects the binding energy, but not the location of the Mg^{2+} -binding site in the pore (Fig. 3e, f).

If substitution of aspartate or glutamate at position 171 influences Mg^{2+} blockade through an electrostatic mechanism, then placement of a cationic amino acid should dramatically reduce the affinity. Figure 4a shows the $I-V$ curve for the mutant channel containing an arginine residue (N171R). As expected, this mutant channel is insensitive to a high concentration of intracellular Mg^{2+} . In addition, the $I-V$ curve rectifies strongly in the absence of Mg^{2+} . To investigate further the Mg^{2+} -independent rectification, we substituted histidine at position 171 (Fig. 4b). The $I-V$ relationship is linear at pH 9.0, but begins to rectify like the arginine mutant channel as the pH is lowered to a range where the cationic form of histidine is favoured. This result

FIG. 1 Intracellular Mg^{2+} blocks the pore of ROMK1 with low affinity. *a*, $I-V$ traces were recorded from excised patches in the presence of 100 mM KCl on both sides of the membrane and different concentrations of intracellular Mg^{2+} . *b*, Fraction of unblocked current is shown as a function of free internal Mg^{2+} concentration at three membrane voltages. Pipette (external) solution contained 100 mM KCl (top) or 20 mM KCl + 80 mM NaCl (bottom). Each point is the mean $\pm$ s.e.m. of 6–12 separate measurements. Curves superimposed on the data are least-squares fits to $I/I_0 = K_i(V)/(K_i(V) + [Mg^{2+}])$. The inhibition constants (in mM) for 20 mM KCl (bottom) are $K_i(0) = 4.4$, $K_i(20) = 2.2$, $K_i(40) = 1.1$, and for 100 mM KCl (top) they are $K_i(30) = 4.2$, $K_i(50) = 2.0$, $K_i(70) = 0.9$. *c*, The $K_i(V)$, determined as for *b*, are plotted as a function of membrane voltage. The lines correspond to least-squares fits to $K_i(V) = K_i(0) \exp(-2\delta FV/RT)$, where K_i is the inhibition constant at 0 mV, δ is the 'electrical distance', V is the membrane voltage and F , R and T have their standard meanings. The fits yield $K_i(0) = 4.5$ mM, $\delta = 0.44$ (open circles; 20 mM KCl) and $K_i(0) = 13.4$ mM, $\delta = 0.50$ (filled circles; 100 mM KCl).

METHODS. Currents from ROMK1 expressed in *Xenopus* oocytes were recorded while ramping membrane voltage from -80 mV to 80 mV over a duration of 0.8 to 1.6 s. Each trace is the leak-subtracted average of 5 records. Pipette solution contained (mM) 100 KCl, 0.3 $CaCl_2$, 1 $MgCl_2$ and 10 HEPES, pH 7.1 . Bath solution contained (mM) 100 KCl, 5 EDTA and 10 HEPES, pH 7.1 (Mg^{2+} -free), or 100 KCl, 0.5 EGTA, 4 Mg -ATP and 10 HEPES, pH 7.1 . Free Mg^{2+} concentration was achieved (maintaining



a fixed concentration of Mg -ATP) by adjusting the total ATP and $MgCl_2$ concentrations using a stability constant of $8.8 \times 10^3 M^{-1}$. Run-down following patch excision was minimal in the presence of internal Mg -ATP or when 5 mM EDTA was present to chelate Mg^{2+} . At the end of each experiment patches were exposed to 2 – 3 mM Mg^{2+} without ATP to induce run-down, and after 3 – 5 min records were taken for leak subtraction. Because very high Mg^{2+} concentrations induce run-down even in the presence of ATP, a small fraction of the current reduction at 10 mM free Mg^{2+} was probably due to run-down. Experiments were at room temperature ($22 \pm 1^\circ C$). All $I-V$ data were filtered at 1 to 5 kHz and stored on a computer disk for subsequent off-line leak subtraction and analysis.

indicates that a positive charge is responsible for the rectifying $I-V$ relationship in the absence of Mg^{2+} . The result also argues that the side chain of residue 171 must be exposed to the ion conduction pathway.

These findings place amino-acid residue 171 close to the Mg^{2+} -binding locus in the ROMK1 K^+ channel. The residue probably does not participate directly in the coordination of Mg^{2+} , but it must be near enough to influence Mg^{2+} binding through an electrostatic mechanism. The Mg^{2+} -independent rectification observed when a cationic residue is placed at position 171 is also explicable in terms of electrostatics. The

curves overlying the data in Fig. 4a–b correspond to a Nernst-Planck electrodiffusion model, and show the effect on the $I-V$ curve of adding positive surface charge to the membrane's internal surface. Increasing the density of positive charge qualitatively mimics the outcome of lowering internal pH in the channel with a histidine residue at 171. This analysis oversimplifies ion conduction in a K^+ channel, but it serves to demonstrate that simple electrostatic principles can account for the effect of mutations at position 171 in ROMK1.

Previous studies have shown that inward rectification in some channels can be attributed at least partly to voltage-dependent

FIG. 2 An acidic residue at position 171 causes strong inward-rectification. *a*, Amino-acid sequence (single-letter code), of the second transmembrane segment (M2) of three inward-rectifier K^+ channels is shown^{11,13,19,20}; the arrow identifies residue 171 in ROMK1. *b*, Sixteen amino acids were substituted at position 171 in ROMK1. Eight mutant channels, indicated by (+), produced currents when expressed in oocytes. The amino acids are listed in order of decreasing side-chain hydrophilicity²⁶. *c*–*f*, The $I-V$ relationship recorded from cell-attached patches is shown for wild type with asparagine at position 171 (WT in *c*), and for mutants containing threonine (*d*; N171T), aspartate (*e*; N171D) and glutamate (*f*; N171E). **METHODS.** Records were obtained and processed as described in Fig. 1 legend. A leak subtraction template was obtained by excising the patch (inside-out) into bath solution without ATP to allow ROMK1 channel run-down. Mutations were produced using PCR mutagenesis (a silent *Nde1* site was created). Mutations were produced and a sequenced cassette (*Nde1*–*Bgl*II; 310–691) containing the mutation was used to replace the corresponding fragment in the wild-type channel. ROMK1 RNA was transcribed *in vitro* from a *NotI*-linearized plasmid (p-SPORT) containing the cDNA using T7 polymerase. *Xenopus* oocytes were prepared as previously described²⁷. Currents were recorded between 2 and 12 days after injection.

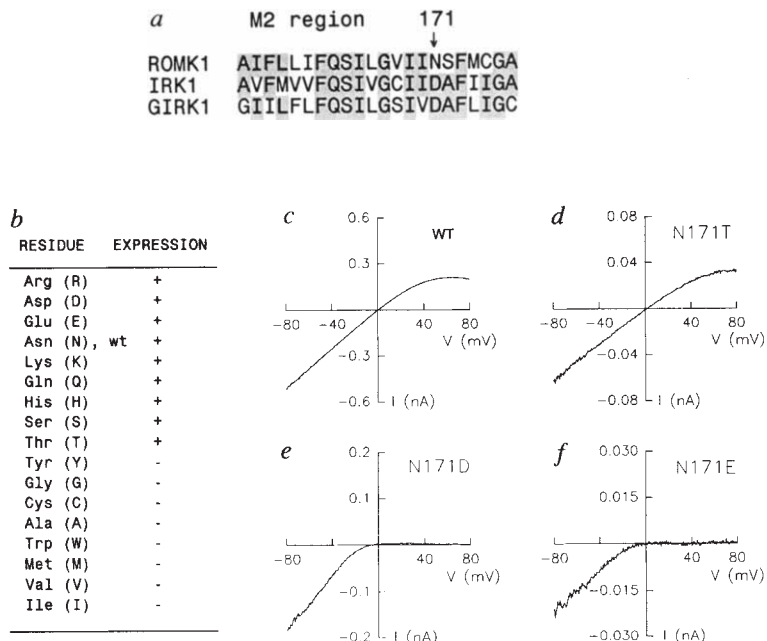
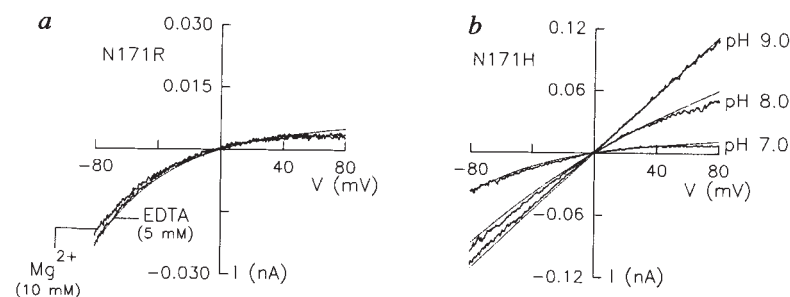
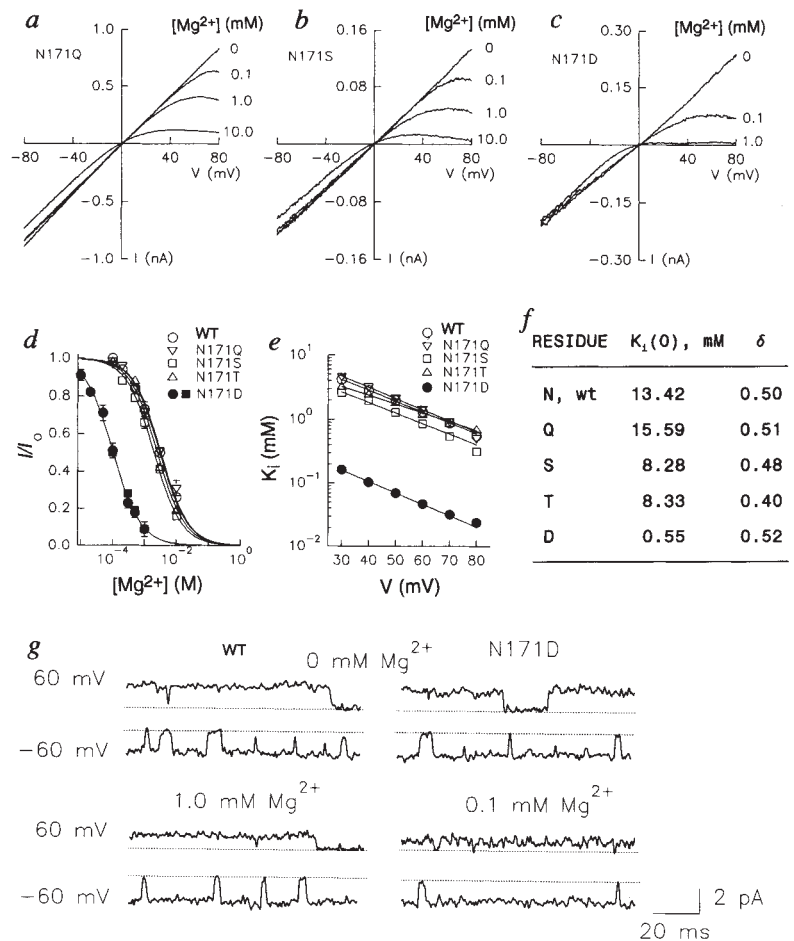


FIG. 3 Mutations at position 171 alter the affinity for internal Mg^{2+} . **a–c**, I - V records from excised inside-out patches containing channels with glutamine (**a**; N171Q), serine (**b**; N171S) and aspartate (**c**; N171D) at position 171 are shown. Free internal Mg^{2+} concentration is given beside each trace. **d**, I/I_0 (mean $\pm$ s.e.m. of 3–12 determinations) measured at 40 mV is plotted as a function of free Mg^{2+} concentration for the wild-type channel and mutant channels N171Q, N171S, N171T and N171D. The curves superimposed on the data are least-squares fits to the equation $I/I_0 = K_i(40)/(K_i(40) + [Mg^{2+}])$. The inhibition constants at 40 mV, $K_i(40)$, are (mM): 3.0 (wild type), 3.2 (N171Q), 2.0 (N171S), 2.5 (N171T) and 0.1 (N171D). To obtain free Mg^{2+} concentrations in the range of 0.01 to 0.05 mM for the N171D mutant, $MgCl_2$ was added directly to an ATP-free bath solution. To minimize run-down in the absence of ATP, 1 mM vanadate was present. These data (filled-circles) agree with those obtained when ATP was present at concentrations of free Mg^{2+} between 0.1 mM to 1 mM (filled squares; Fig. 1 legend). **e**, The inhibition constant $K_i(V)$ is plotted as a function of membrane voltage for wild-type and mutant channels. $K_i(V)$ were determined as in **d**. Lines correspond to a least-squares fit to the equation $K_i(V) = K_i(0) \exp(-2\delta FV/RT)$. **f**, The zero-voltage inhibition constant $K_i(0)$ and electrical distance δ for data in **e** are tabulated. **g**, Single wild-type and N171D mutant channels at -60 mV and 60 mV with and without Mg^{2+} . Data were filtered at 1 kHz.



gating^{8,21–24} and that other regions of inward rectifier K^+ channels may also contribute to Mg^{2+} sensitivity²⁵. Multiple determinants apparently underlie inward rectification. We have shown using site-directed mutagenesis that a single amino acid in the second (putative) membrane-spanning segment of ROMK1 strongly influences Mg^{2+} affinity through an electrostatic mechanism. We anticipate that the amino acid at the equivalent location in other inward rectifiers will also influence their Mg^{2+} sensitivity. □

FIG. 4 Basic-amino-acid substitutions point to an electrostatic interaction between residue 171 and cations in the pore. **a**, The I - V relationship for channels with an arginine residue (N171R) was recorded from an inside-out patch in the absence of Mg^{2+} (5 mM EDTA) and in the presence of 10 mM free internal Mg^{2+} . The smooth curve corresponds to the Goldman current equation with an internal positive surface charge: $I = g(V_h + \phi)[c] \exp(-F\phi/RT) \{1 - \exp(-FV_h/RT)\} / \{1 - \exp(-F(V_h + \phi)/RT)\}$, where internal surface potential $\phi = (2RT/F) \ln \{ (136\sigma/\sqrt{[c]}) + \sqrt{(136\sigma/(\sqrt{[c]})^2 + 1)} \}$, (external surface potential = 0), V_h is the voltage clamp potential, σ is surface charge density (0.007 unit charges per $\text{\AA}^2$) and $[c]$ is the concentration of KCl (0.1 M) resulting in $\phi = 91$ mV (refs 28, 29). **g** is a scaling factor proportional to conductance. **b**, The I - V relationships measured at 3 different internal pH values are shown for channels with a histidine at residue 171. Internal solution contained 5 mM EDTA and no added Mg^{2+} . The curves on the data correspond to the same model with $\sigma = 0$ (pH 9.0), 0.0014 (pH 8.0) and 0.0047 (pH 7.0) unit charges per $\text{\AA}^2$ (corresponding to surface potentials of 0, 29 and 73 mV, respectively).

Received 13 May; accepted 15 July 1994.

- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA, 1991).
- Trautwein, W. & Dudel, J. *Pflügers Arch.* **266**, 324–334 (1958).
- Noma, A., Peper, K. & Trautwein, W. *Pflügers Arch.* **400**, 424–431 (1983).
- Sakmann, B., Noma, A. & Trautwein, W. *Nature* **303**, 250–253 (1983).
- Seojima, M. & Noma, A. *Pflügers Arch.* **400**, 424–431 (1984).
- Ashcroft, F. M. & Rorsman, P. *Biochem. Soc. Trans.* **18**, 109–111 (1990).
- Horie, M., Isisawa, H. & Noma, A. *J. Physiol.* **387**, 251–272 (1987).
- Matsuda, H., Saigusa, A. & Isisawa, H. *Nature* **325**, 156–159 (1987).
- Vandenberg, C. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2560–2564 (1987).
- Matsuda, H. *J. Physiol.* **435**, 83–99 (1991).
- Ho, K. et al. *Nature* **362**, 31–38 (1993).

12. Nichols, C. G., Ho, K. & Hebert, S. J. *Physiol.* **476**, 399–409 (1994).
13. Kubo, Y., Baldwin, T. J., Yan, Y. N. & Jan, L. Y. *Nature* **362**, 127–132 (1993).
14. Miller, C. *Science* **252**, 1092–1096 (1991).
15. Choi, K. L., Mossman, C., Aube, J. & Yellen, G. *Neuron* **10**, 533–541 (1993).
16. Kirsch, G. E., Shieh, C. C., Drewe, J. A., Vener, D. F. & Brown, A. M. *Neuron* **11**, 503–512 (1993).
17. Slesinger, P. A., Jan, Y. N. & Jan, L. Y. *Neuron* **11**, 739–749 (1993).
18. Lopez, G. A., Jan, Y. N. & Jan, L. Y. *Nature* **367**, 179–182 (1994).
19. Kubo, Y., Reuveny, E., Slesinger, P. A., Jan, Y. N. & Jan, L. Y. *Nature* **364**, 802–806 (1993).
20. Dascal, N. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10235–10239 (1993).
21. Stanfield, P. R. et al. *J. Physiol.* **475**, 1–7 (1994).
22. Kurachi, Y. *J. Physiol.* **366**, 365–385 (1985).
23. Ishihara, K., Mitsuiye, T., Noma, A. & Takano, T. *J. Physiol.* **419**, 297–320 (1989).
24. Silver, M. R. & DeCoursey, T. E. *J. gen. Physiol.* **96**, 109–133 (1990).
25. Tagliatela, M., Wibb, B. A., Caporaso, R. & Brown, A. M. *Science* **264**, 844–847 (1994).
26. Creighton, T. E. *Proteins: Structures and Molecular Properties* 2nd edn (Freeman, New York, 1993).
27. MacKinnon, R., Reinhart, P. H. & White, M. M. *Neuron* **1**, 997–1001 (1988).
28. Goldman, D. E. *J. gen. Physiol.* **27**, 37–60 (1943).
29. Grahame, D. *Chem. Rev.* **41**, 441–501 (1947).

ACKNOWLEDGEMENTS. We thank K. Ho and S. Hebert for ROMK1 cDNA, C. McNicholas for technical advice and sharing unpublished results, K. Swartz for discussion, and B. Bean for critically reviewing the manuscript. Z. Lu was a Harvard Mahoney Fellow. This research is supported by a grant from the NIH.

Gating of inwardly rectifying K^+ channels localized to a single negatively charged residue

Barbara A. Wible*, Maurizio Tagliatela*†, Eckhard Ficker* & Arthur M. Brown*‡

* Department of Molecular Physiology and Biophysics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

† Department of Human Communication Sciences, Section of Pharmacology, 2nd School of Medicine, University of Naples 'Federico II', 5 via S. Pansini, 80121 Naples, Italy

INWARDLY rectifying K^+ channels (IRKs) conduct current preferentially in the inward direction. This inward rectification has two components: voltage-dependent blockade by intracellular Mg^{2+} (Mg_i^{2+})^{1–7} and intrinsic gating^{8,9}. Two members of this channel family, IRK1 (ref. 10) and ROMK1 (ref. 11), differ markedly in affinity for Mg_i^{2+} (ref. 12). We found that IRK1 and ROMK1 differ in voltage-dependent gating and searched for the gating structure by large-scale and site-directed mutagenesis. We found that a single amino-acid change within the putative transmembrane domain M2, aspartate (D) in IRK1 to the corresponding asparagine (N) in ROMK1, controls the gating phenotype. Mutation D172N in IRK1 produced ROMK1-like gating whereas the reverse mutation in ROMK1—N171D—produced IRK1-like gating. Thus, a single negatively charged residue seems to be a crucial determinant of gating.

Currents from IRK1 increase over several milliseconds to a steady level on hyperpolarization whereas ROMK1 currents show no comparable time dependence (Fig. 1). To localize the structural domains responsible for gating, we constructed chimaeras between IRK1 and ROMK1 channels. Exchange of the H5 domain¹² had no effect on gating or rectification as shown by chimaeras CHM A and REV CHM A. However, substitution of the entire membranous core (M1–H5–M2) of IRK1 for the corresponding ROMK1 sequence (CHM B) replaced the slower gating of IRK1 with the time-independent gating characteristic of ROMK1 but did not affect Mg_i^{2+} blockade¹². Therefore, residues in the M1 and M2 domains were likely gating determinants.

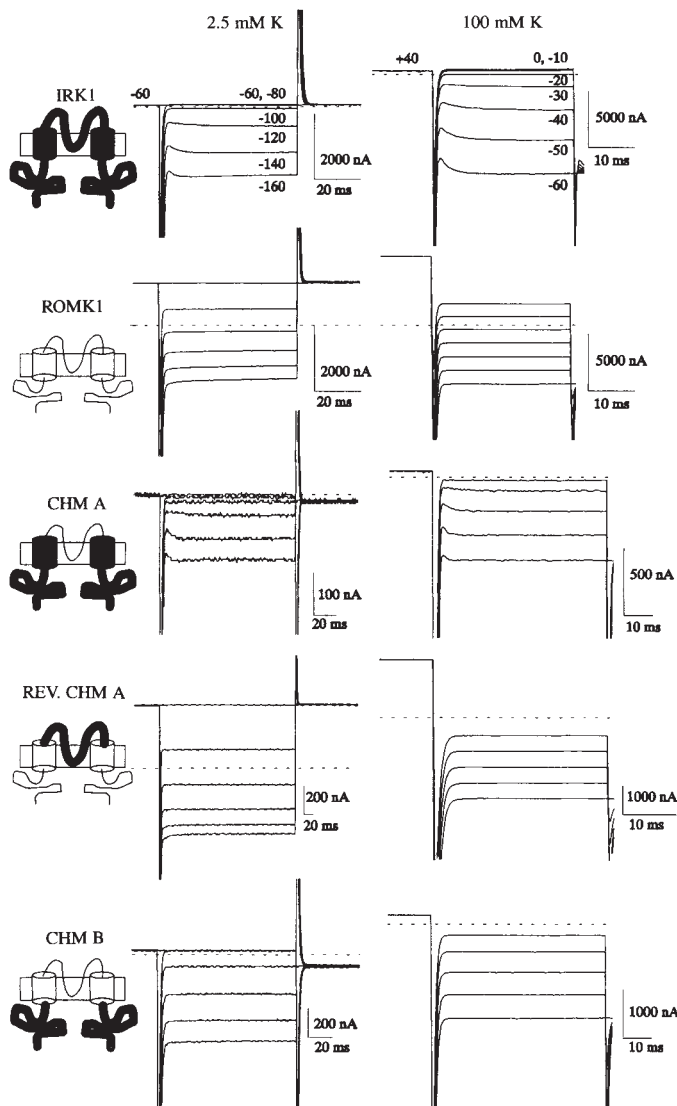


FIG. 1 Localization of the gating domain in IRKs. Macroscopic currents were recorded using a two-microelectrode voltage clamp of *Xenopus* oocytes²⁰ injected with cRNAs encoding ROMK1, IRK1 and the chimaeric constructs CHM A, REV.CHM A and CHM B. Left panels: holding potential -60 mV; 60 – 200 -ms steps from -160 mV to -60 mV in 20 -mV increments; return potential: -60 mV (except in CHM A and CHM B, where it was -100 mV). Right panels: holding potential -60 mV; depolarizing 100 -ms prepulse to $+40$ mV, followed by returns from -60 mV to 0 mV in 10 -mV increments (IRK1 and ROMK1), or from -60 mV to -20 mV in 10 -mV increments (CHM A, REV.CHM A and CHM B); return potential: -60 mV (except in CHM A and CHM B, where it was -100 mV). Currents at the end of the depolarizing prepulse and during hyperpolarizing pulses are shown. Dashed lines indicate zero current. Part of the capacitive artefacts have been clipped. At low $[K^+]_o$ and very negative potentials, ROMK1 and REV CHM A currents are reduced, possibly as a result of depletion and a reduction in P_o (ref. 11).

METHODS. The chimaeric cDNA constructs between ROMK1 and IRK1 were previously described¹². Single point mutations in IRK1 and ROMK1 were made by polymerase chain reaction (PCR) overlap extension²¹, and subcloned into the A'-pCRLII vector for expression in *Xenopus* oocytes as previously described¹². Bath solutions were (in mM): low K^+ -MES (left panels) (2.5 KOH, 120 N-methyl-D-glucamine, 122.5 2-[N-morpholino]ethanesulphonic acid (MES), 2 Mg(OH)₂ and 10 HEPES, pH 7.3); intermediate K^+ -MES (22.5 KOH, 100 N-methyl-D-glucamine, 122.5 MES, 2 Mg(OH)₂ and 10 HEPES, pH 7.3); or high K^+ -MES (right panels) (100 KOH, 22.5 N-methyl-D-glucamine, 122.5 MES, 2 Mg(OH)₂ and 10 HEPES, pH 7.4). The pClamp software (Axon Instruments) was used for the generation of the voltage pulse protocols and for data acquisition.

‡ To whom correspondence should be addressed.

As gating is voltage-dependent, we searched for differences in charged residues in domains M1 and M2 (Fig. 2a). Only one difference is present: in IRK1, residue 172 is a D whereas the corresponding residue 171 in ROMK1 is an N. Single point mutations were made at these positions. For IRK1, the D/N mutant resulted in currents with virtually instantaneous activation kinetics characteristic of ROMK1 gating (Fig. 2b). A recent report has also implicated this position in IRK1 gating¹³. Conversely, the N/D mutation in ROMK1 expressed currents with the slower onset of activation characteristic of IRK1 channels (Fig. 2b). Thus, a single amino acid controls the gating phenotype of these two IRKs.

Outward currents were negligible for IRK1 (Fig. 3a) and very small for IRK1 D/N (Fig. 3b) and ROMK1 N/D (Fig. 3c). In IRK1 and ROMK1, N/D inward currents elicited on hyperpolarization from +40 mV ($E_K = 0$ mV) activated with time constants (τ) between 0.5 and 10 ms. The time-course of activation was monoexponential and voltage-dependent, the rates increasing with more hyperpolarizing pulses (Fig. 3d, e). By contrast, activation of the mutant IRK1 D/N was virtually instantaneous (Fig. 3b). The voltage-dependence of the activation rates for IRK1 and ROMK1 N/D was similar (Fig. 3e). The relationship between activation τ values and voltage was shifted to the right with increased $[K^+]_0$ (Fig. 3d)^{14,15}. In IRK1, a change in E_K of

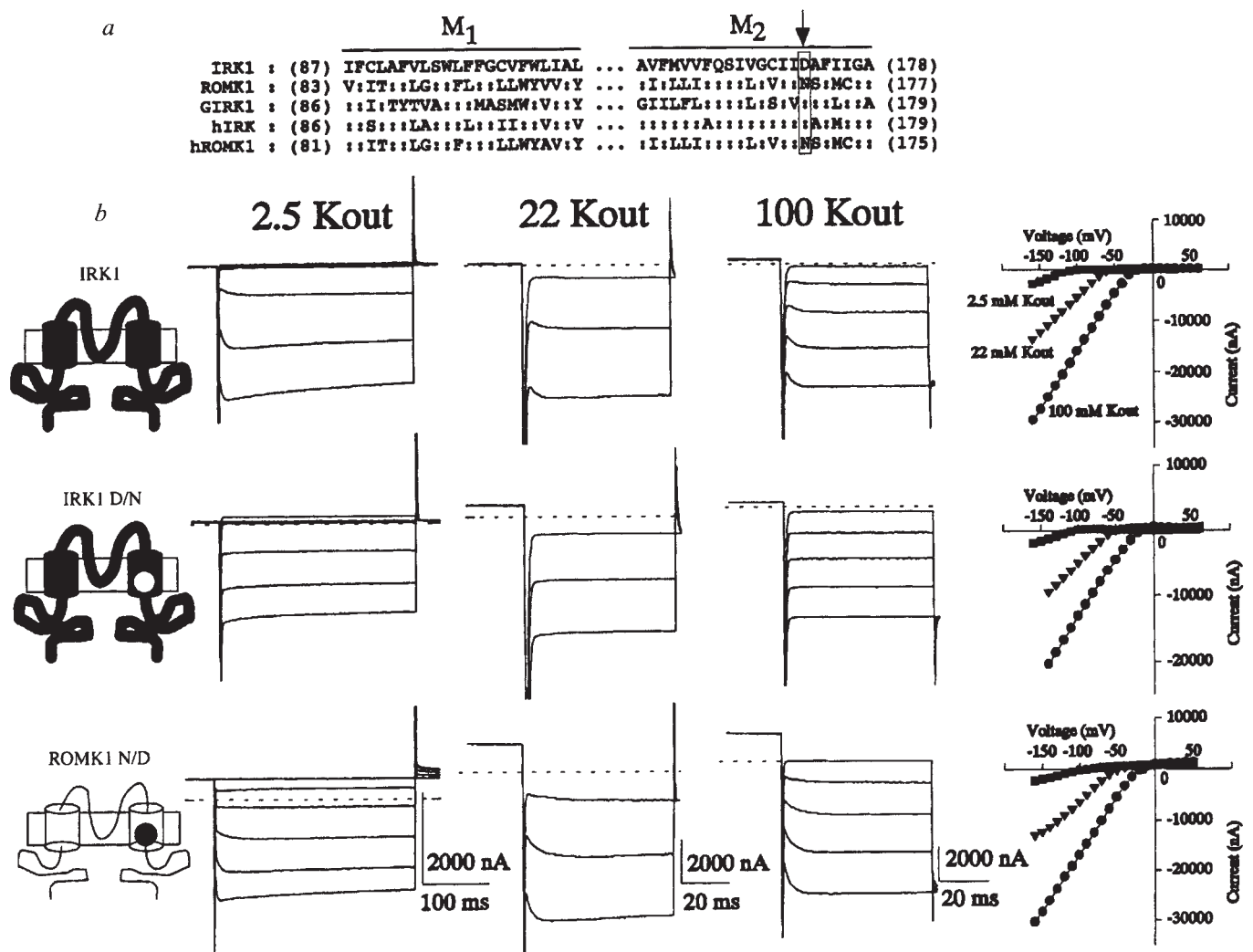
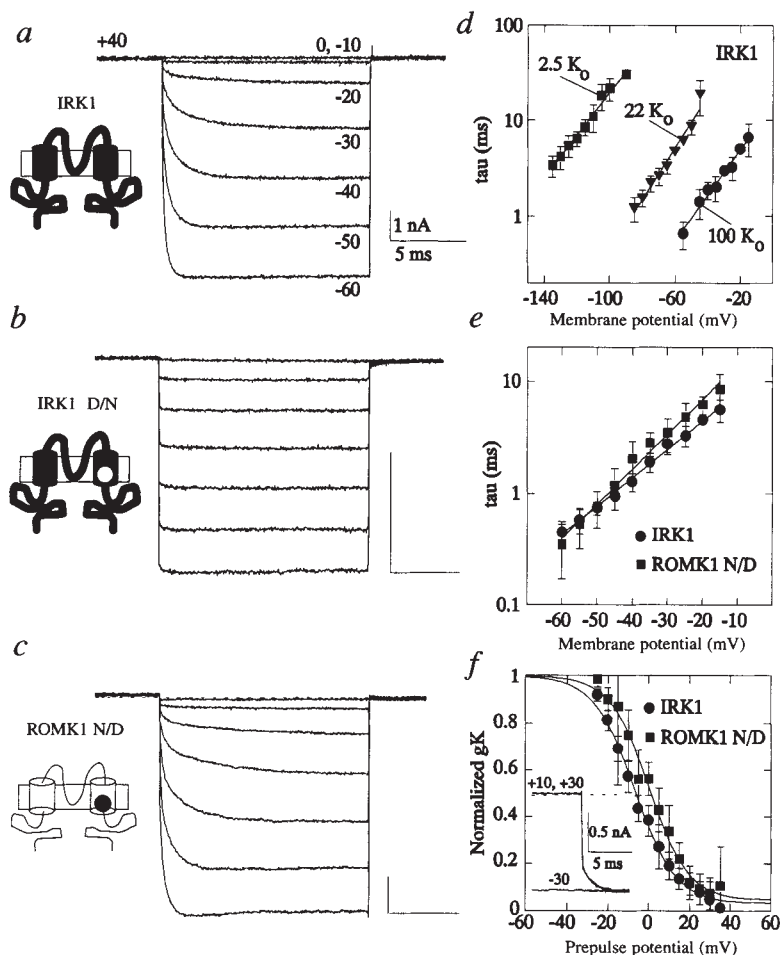


FIG. 2 Localization of the gating of IRKs to a single residue. *a*, Amino acid sequences of the putative transmembrane domains M1 and M2 of IRKs are aligned^{10,11,22}. Colons indicate amino acid identities to the IRK1 sequence. hIRK is a human heart IRK we have cloned (B.A.W. et al., manuscript submitted) which shows strong inward rectification and marked time-dependent activation. hROMK1, an IRK cloned from human kidney²³, has functional properties analogous to those of ROMK1. *b*, Currents recorded with the two-microelectrode voltage clamp from *Xenopus* oocytes injected with cRNAs encoding IRK1, IRK1 D/N and ROMK1 N/D. Left current traces: holding potential -60 mV; 100-300-ms steps from -160 mV to -60 mV in 20-mV increments; post-return potential: -60 mV. Middle current traces: holding potential -60 mV; depolarizing 100-ms prepulse to +40 mV, followed by

returns from -100 mV to -60 mV in 20-mV steps; post-return potential: -60 mV. Right current traces: holding potential -60 mV; depolarizing 100-ms prepulse to +40 mV, followed by returns from -60 mV to -20 mV in 10-mV steps; post-return potential: -60 mV. Dashed lines indicate zero current. Bath solutions were: low K^+ -MES (left current traces), intermediate K^+ -MES (middle current traces) or high K^+ -MES (right current traces) (see Fig. 1 legend). Part of the capacitive artefacts have been clipped. In the right panels, the relationship of the steady-state current (measured at the end of the pulses) to voltage (from -160 mV to +60 mV) is shown at three different $[K^+]_0$: low K^+ -MES (■, 2.5 mM); intermediate K^+ -MES (▼, 22 mM); or high K^+ -MES (●, 100 mM).

FIG. 3 Gating kinetics of IRK1, IRK1 D/N, and ROMK1 N/D. Cell-attached macropatch recordings from *Xenopus* oocytes injected with IRK1 (a) IRK1 D/N (b) and ROMK1 N/D (c) cRNAs. Holding potential: 0 mV; depolarizing 100-ms prepulse to +40 mV, test pulses from -60 mV to 0 mV in 10-mV increments. Dashed lines indicate zero current. Pipette solution: K⁺-Ringer, bath solution: iso-K⁺. Leak and capacitive currents have been subtracted by a P/6 procedure from +40 mV. **d**, Voltage-dependence of time constants measured from mono-exponential fits to test currents obtained with a two-microelectrode voltage clamp as described in Figs 1 and 2. For comparison with macropatch recordings, the data were corrected for a -15-mV offset resulting from a liquid junction potential. **e**, Time constants measured from monoexponential fits to IRK1 (a, ●), or ROMK1 N/D (c, ■). τ values of 500 μ s or less could not be evaluated. **f**, Voltage-dependence of the normalized K⁺ conductance (g_{K+}) for IRK1 (●) and ROMK1 N/D (■). g_{K+} values were obtained with the following protocol: holding potential, 0 mV; 100-ms prepulse from -35 mV to +25 mV, 5-mV increments; test pulse to -30 mV. The insert shows traces obtained with prepulses to -30 mV, +10 mV and +30 mV. Normalized current amplitudes were plotted as a function of prepulse potential and fitted by Boltzmann equations of the following form: $g_{K+}/g_{K+}^{\infty} = 1/(1 + \exp((V - V_{1/2})/kT))$, where V is the prepulse potential, $V_{1/2}$ is the half-activation potential, z is the effective valence. e , k and T are thermodynamic constants. E_K values were -0.47 ± 1.05 mV and -0.96 ± 1.38 mV for IRK1 ($n=5$) and ROMK1 N171D ($n=5$), respectively.

METHODS. For macropatch recordings²⁴ the pipette solution was K⁺-Ringer (in mM: 100 KCl, 2 MgCl₂, 10 HEPES, pH 7.3), while the bath solution was iso-K⁺ (in mM: 100 KCl, 10 EDTA, 10 HEPES, pH 7.3). When necessary, MgCl₂ or Mg-ATP was added to the bath solution, and



EGTA was substituted for EDTA to give the calculated free Mg²⁺ concentrations²⁵. Data were filtered at 8 kHz and digitized at 33 kHz.

-29.5 ± 1.2 mV ($n=4$) and -84.5 ± 5.2 mV ($n=4$) resulted in shifts of -38.8 ± 1.5 mV ($n=4$) and -105.0 ± 2.1 mV ($n=4$), respectively. Similar shifts were seen for ROMK1 N/D: for changes in E_K of -29.2 ± 1.3 mV ($n=5$) and -76.4 ± 3.1 mV ($n=5$), the voltage-dependence of the τ values was shifted by -37.5 ± 1.5 mV ($n=4$) and -88.9 ± 1.5 mV ($n=5$), respectively. The steady-state dependence of the conductance on voltage was also similar for IRK1 and the ROMK1 N/D mutant (Fig. 3f). For IRK1 and ROMK1 N/D, respectively, the half-activation potentials were -0.39 ± 4.7 mV ($n=5$) and -6.73 ± 2.45 mV ($n=5$), and the effective valences were 2.74 ± 0.24 ($n=5$) and 2.80 ± 0.51 ($n=5$). Taken together, these results confirm that a single N/D mutation in M2 confers IRK1-like gating to ROMK1.

Despite the appearance of instantaneous gating, macroscopic currents of IRK1 D/N retained the marked inward rectification of IRK1 (Fig. 2b), suggesting that Mg²⁺ block was unaffected by the mutation. Mg²⁺ produced a fast blockade of both IRK1 and IRK1 D/N single channels causing, at a free concentration of 0.288 mM, an increase in open channel noise (Fig. 4a). For IRK1 D/N, the dissociation binding constant (K_d) for Mg²⁺ blockade at +40 mV was 30.4 ± 8 μ M ($n=4$), a value similar to the K_d of 17 μ M for IRK1¹². To measure the K_d values over a range of potentials, we recorded ramp currents produced by single channels in cell-attached patches and excised patches exposed to zero Mg²⁺ or 0.288 mM Mg²⁺. The ramp currents were similar for IRK1 and IRK1 D/N (Fig. 4b). The fraction

of the electrical field felt by Mg²⁺, δ , was 0.59 ± 0.07 ($n=3$) for IRK1 and 0.57 ± 0.11 ($n=3$) for IRK1 D/N (see Fig. 4 legend for calculation). Thus the voltage dependence of Mg²⁺ block was not significantly changed by the mutation.

The single-channel recordings also showed that the mutant IRK1 D/N has a conductance for K⁺ similar to IRK1 (18.8 ± 0.4 pS, $n=9$, and 21.4 ± 0.4 pS, $n=11$, respectively). Likewise, ROMK1 N/D and ROMK1 had comparable single-channel K⁺ conductances (33.7 ± 0.4 pS, $n=10$, and 31.9 ± 0.4 pS, $n=7$, respectively). The affinity for Mg²⁺ was increased, however, in the ROMK1 N/D mutant ($K_d = 220$ μ M at +40 mV, $n=2$) compared with ROMK1 ($K_d = 1.7$ mM at +40 mV; ref. 12).

The most direct way to separate gating and Mg²⁺ blockade is to study kinetics in the absence of Mg²⁺. The gating characteristics of IRK1 and IRK1 D/N inward currents described in the cell-attached configuration were not altered in Mg²⁺-free solutions (Fig. 4b, left). Removal of Mg²⁺ produced much larger outward currents which deactivated markedly in IRK1 as compared with IRK1 D/N (Fig. 4b, right). These results were equally evident in single-channel patches excised from IRK1 and IRK1 D/N (Fig. 4a). For IRK1, average single-channel P_o at +40 mV decreased in a time-dependent manner in the complete absence of Mg²⁺ (data not shown)¹⁶. The time-dependent decrease in P_o , however, was absent in IRK1 D/N.

The persistence of high Mg²⁺ affinity despite the introduction of virtually instantaneous gating explains the strong rectification

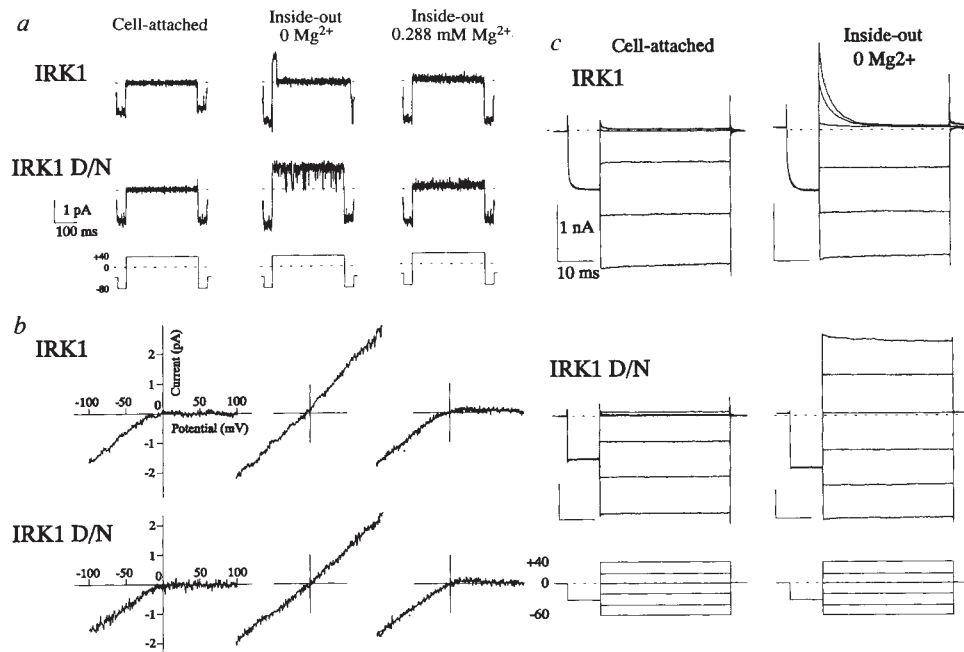


FIG. 4 Gating and Mg^{2+} block. **a**, Single-channel currents from IRK1 or IRK1 D/N channels. Patches were recorded in the cell-attached configuration and upon excision into an internal solution containing either zero or 0.288 mM free Mg^{2+} in the inside-out configuration. Pipette solution: K^+ -Ringer; bath solution: iso- K^+ . The K_d for Mg^{2+} was calculated from single-channel currents obtained from inside-out patches. 32–64 sweeps were recorded at 0, 0.0052, 0.052, 0.288 and 1.7 mM free Mg^{2+} . The integrated current values were expressed as a function of $[Mg^{2+}]$ and fitted to the following binding isotherm: $y = \max/(1 + X/K_d)^n$, where X is the Mg^{2+} concentration and n the Hill coefficient¹². The fitted values for n were between 0.8 and 1.0. **b**, Single-channel ramp currents recorded from IRK1 or IRK1 D/N in cell-attached patches and after excision into zero Mg^{2+} and 0.288 mM Mg^{2+} . Each record is the average of three to eight (0.625 mV ms^{-1}) sweeps analysed as described²⁶. The ramp currents in the inside-out patches exposed to Mg^{2+} were fitted to the equation: $i = g_i(V - V_K)/K_{d(V)}$ where i is single-channel current, g_i is single channel conductance, V is membrane potential, V_K is potassium equilibrium potential and K_d the 1:1 dissociation binding constant for Mg^{2+} to the channel. $K_{d(V)} = K_{d(0)} \exp(z\delta VF/RT)$ ²⁷ where δ is the voltage-dependence of binding and z , F , R and T have their usual thermodynamic meaning. The K_d values at +40 mV obtained from the ramps

were 33 ± 12 ($n=3$) and 43 ± 10 ($n=2$) μM for IRK1 and IRK1 D/N, respectively. These values did not differ significantly from those obtained with constant voltage steps. **c**, Macroscopic currents recorded from IRK1 or IRK1 D/N channels in the cell-attached configuration (left panels) and upon patch excision into a zero Mg^{2+} solution (right panels). Holding potential: 0 mV; hyperpolarizing prepulse potential: -30 mV ; depolarizing pulses from -60 to $+40 \text{ mV}$ steps; post-return potential 0 mV. Dashed lines indicate zero current. Pipette solution: K^+ -Ringer, bath solution: iso- K^+ . Current traces are shown without subtraction of leak and capacitive currents. The gating properties in both single channels and macropatch inside-out recordings of IRK1 displayed slowly changing kinetics with a rather slow increase in the time constant of the outward current decay. These phenomena, which have also been described in cardiac myocytes^{1–7}, suggest that either a soluble gating component has been washed away, or that the inside-out configuration promoted some modification of the channel itself²⁸. **METHODS**. Single-channel currents were recorded at room temperature ($22\text{--}24^\circ\text{C}$) using fire-polished and Sylgard-coated micropipettes of 2–5 M Ω resistance. The pipette and bath solutions were identical to those used for studying macropatches. Data were low-pass filtered at 1 kHz before digitization at 5–10 kHz.

of IRK1 D/N currents and suggests a dissociation between gating and ion blockade. This dissociation is less clear cut for ROMK1 N/D in which a large increase in Mg^{2+} affinity accompanies IRK1-like gating. In this case, both gating and Mg^{2+} block seem to be determined largely by a single amino acid. The control of gating by a single negatively charged residue in IRKs is in marked contrast to voltage-gated ion channels in which four to eight positively charged residues in the S4 domain are implicated in the voltage-dependent activation process^{17–19}. □

Received 31 May; accepted 18 August 1994.

- Matsuda, H., Saigusa, A. & Irisawa, H. *Nature* **325**, 156–159 (1987).
- Vandenberg, C. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2560–2564 (1987).
- Matsuda, H. *Rev. Physiol.* **53**, 289–298 (1991).
- Matsuda, H. in *Mg²⁺ and Excitable Membranes* (eds Strata, P. & Carbone, E.) 51–70 (Springer, Berlin/Heidelberg, 1991).
- Anumonwo, J. M. B., Freeman, L. C., Kwok, W. M. & Kass, R. S. *Cardiovasc. Drug Rev.* **9**(3), 299–316 (1991).
- Shimoni, Y., Clark, R. B. & Giles, W. R. *J. Physiol., Lond.* **448**, 709–727 (1992).
- Hille, B. in *Ionic Channels of Excitable Membranes*, Vol. 2, 83–114 (Sinauer Associates, Sunderland, 1992).
- Ciani, S., Krasne, S. & Hagiwara, S. *Biophys. J.* **30**, 199–204 (1980).

- Ishihara, K., Mitsuiye, T., Noma, A. & Takano, M. *J. Physiol., Lond.* **419**, 297–320 (1989).
- Kubo, Y., Baldwin, T. J., Jan, Y. N. & Jan, L. Y. *Nature* **362**, 127–133 (1993).
- Ho, K. et al. *Nature* **362**, 31–38 (1993).
- Taglialatela, M., Wible, B. A., Caporaso, R. & Brown, A. M. *Science* **264**, 844–847 (1994).
- Stanfield, P. R. et al. *J. Physiol., Lond.* **478**, 1–6 (1994).
- Leech, C. A. & Stanfield, P. R. *J. Physiol., Lond.* **319**, 295–309 (1981).
- Pennefather, P., Oliva, C. & Mulrine, N. *Biophys. J.* **61**, 448–462 (1992).
- Stanfield, P. R. et al. *J. Physiol., Lond.* **475**(1), 1–7 (1994).
- Jan, L. Y. & Jan, Y. N. *Nature* **345**, 672 (1990).
- Catterall, W. A. *Trends neurochem. Sci.* **16**, 500–506 (1993).
- Aldrich, R. *Nature* **362**, 107–108 (1993).
- Taglialatela, M. et al. *Pflügers Arch. ges. Physiol.* **423**, 104–112 (1993).
- Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. *Gene* **77**, 51–59 (1989).
- Kubo, Y., Reuveny, E., Slesinger, P. A., Jan, Y. N. & Jan, L. Y. *Nature* **364**, 802–806 (1993).
- Kugler, J. L., Yano, H., Tokuyama, Y. J., Nelson, D. J. & Philipson, L. *Biophys. J.* **66**, A427 (1994).
- Hilgemann, D. W. *Pflügers Arch. ges. Physiol.* **415**, 247–249 (1989).
- Fabiato, A. & Fabiato, F. *J. Physiol. (Paris)* **75**, 463–505 (1979).
- Kirsch, G. E. et al. *Biophys. J.* **62**, 136–144 (1992).
- Woodhull, A. M. *J. gen. Physiol.* **62**, 687–708 (1973).
- Ruppertsberg, J. P. et al. *Nature* **352**, 711–714 (1991).

ACKNOWLEDGEMENTS. We thank L. Jan for providing the IRK1 clone; W.-Q. Dong, D. Jackson and R. Cockrell for injecting and handling oocytes; T. Affini, C. Zhu, S. Dou and S. Pan for molecular biology support; R. Caporaso for help throughout the project; A. E. Lacerda for help with the models for Mg^{2+} blockade; and G. E. Kirsch and A. E. Lacerda for discussion. This work was supported in part by NIH grants to A.M.B. and a grant from A. von Humboldt Foundation to E.F.

Peptides determine the lifespan of MHC class II molecules in the antigen-presenting cell

Christopher A. Nelson, Shirley J. Petzold & Emil R. Unanue

Center for Immunology and Department of Pathology, Washington University School of Medicine, St Louis, Missouri 63110, USA

ALTHOUGH many peptides are generated during the intracellular processing of protein antigens, only a few are selected for recognition by the immune system¹⁻⁵. The immunodominant epitope of hen egg white lysozyme (HEL) for H-2^k mice is contained in a tryptic fragment of amino-acid residues 46–61 (refs 6, 7). The core of this T-cell epitope, from amino acids 52 to 61 (DYGILQINSR), contains those residues required for binding to the class II molecule I-A^k (ref. 7). Most of the naturally processed fragments recovered from I-A^k-bearing antigen-presenting cells (APCs) cultured with HEL contained this 52–61 core sequence, presented as a nested set of peptides with extensions at both the amino and carboxyl termini⁸. We now compare the handling by APCs of peptides containing HEL 52–61 to establish whether there is an advantage for the APC in selecting extended peptides: different complexes between peptides and major histocompatibility complex (MHC) molecules varied greatly in the amount of time associated with the APC, and in their immunogenic strength. This difference in persistence is one of the factors contributing to the selection and immune recognition of peptide–MHC complexes by T cells.

The half-lives of peptide–I-A^k complexes were measured in APCs pulsed with radiolabelled peptides. After binding, the 46–61 peptide remained associated with I-A^k without much loss, yielding a half-life of greater than 25 h in the three B-cell lymphomas tested, CH27 (ref. 9), TA3 (ref. 10), and M12C3.F6 (ref. 11) (abbreviated to C3.F6) (Fig. 1). A similar long-lived association was observed for the 48–62 peptide, which represents the most abundant fragment recovered from APCs after culture in HEL⁸ (Table 1). By contrast, the 52–61 core peptide bound to I-A^k but was lost with a half-life of about 7 h in all three APC lines (Fig. 1, Table 1). These differences in half-lives between the 46–61 and the 52–61 peptides were also reflected in the number of biologically relevant complexes at the APC surface (Table 2). Previous studies of half-lives of peptides bound to HLA-DR molecules of B-lymphoblastoid cells had indicated that the peptide–MHC complex was practically irreversible¹². These results agree with the finding of long-lived complexes but indicate that differences in half-life exist among peptides.

The half-lives of the labelled 46–61 and 52–61 peptides were unchanged by incubation in excess unlabelled HEL 46–61 or HEL, which should discourage the reassociation of any released labelled peptide with I-A^k (Fig. 1b). Therefore, the persistence of the 46–61 or 48–62 peptides cannot be explained by exchange of this peptide between class II molecules; and the at least 25-h half-life must represent that of the peptide-associated class II molecule. As the half-life of leucine-labelled I-A^k molecules in these B-cell lymphomas ranged over 14–18 h (medium only in Fig. 2), the 46–61 peptide must extend the life of the class II molecule to which it is bound. Clearly the loss of the 52–61 peptide is not accelerated by excess unlabelled peptide, either through a mechanism of ligand-induced dissociation¹³ or competitive exchange¹⁴. The short half-life of the 52–61 peptide represents either the dissociation of this peptide from the class II molecules of the APC or the selective degradation of the class II molecules to which this peptide is bound. Notably, that

although the long 46–61 and the short 52–61 peptides differ in affinity (30-fold)¹⁵, this difference is not reflected in dissociation times measured *in vitro*. A low dissociation rate was also observed for the release of these peptides from complexes in cell lysates, done either at 21 °C or 37 °C (Fig. 1c).

The half-life of the I-A^k molecule in C3.F6 cells was altered by replacing the endogenously presented peptides with peptides generated from HEL (Fig. 2). C3.F6 cells, metabolically labelled with leucine, were cultured in HEL, and after 4 h the cells were washed and placed back in culture with medium alone. The average half-life of the I-A^k molecules for these cells was prolonged from 12.6 h to 23 h (Fig. 2; total I-A^k). For the experiment described in Fig. 2, we also compared half-lives of the stable and unstable I-A^k molecules as determined by SDS–polyacrylamide gel electrophoresis (SDS–PAGE)^{16–21}. The half-lives of the stable molecules from APCs grown in medium were longer than those of the unstable molecules (20.3 h compared with 10.2 h) (Fig. 2; medium only). This difference between the stable and unstable half-lives was even more pronounced for cells incubated in HEL (43.0 h and 10.9 h) (Fig. 2; media + HEL). The addition of HEL did not significantly alter the half-life of the unstable molecules. Therefore, the increase in half-life of the total I-A^k after loading with HEL was due directly to an increase in the relative amount of the SDS-stable molecules, which was evident after culture with HEL.

Table 1 examines the half-life in APCs and the SDS–PAGE stability properties of different MHC–peptide 52–61 complexes. Peptide 51–61 (threonine is the natural residue at position 51) bound to I-A^k and induced a mixed pattern in SDS–PAGE. The half-life in APC was similar to that of the average I-A^k molecules. When the threonine residue at 51 was changed to alanine, the complexes produced were unstable and yielded a short half-life. Substitution at residue 51 with lysine, arginine or leucine all resulted in half-lives in the cell close to the mean half-life of I-A^k. But substitution at this position with a negatively charged glutamic-acid residue resulted in peptides that formed mostly

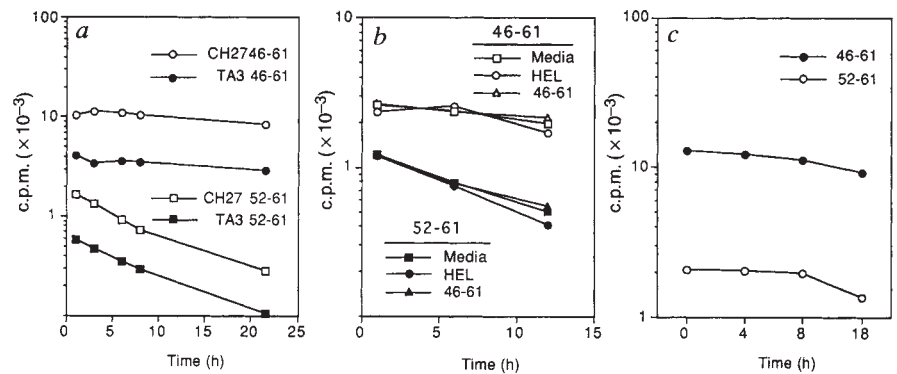
TABLE 1 Time of persistence of I-A^k-labelled peptides in APCs

Peptide	T _{1/2} (h)	n	% Stable
52–61 (DYGILQINSR)	6.8(±2.1)	10	12
52–62 (52–61–W)	8.0	1	16
A-52–61	9.6 (±4.4)	4	24
L-52–61	11.5	1	10
R-52–61	12.8 (±4.6)	4	12
D-52–61	13.2 (1.7)	3	60
K-52–61	13.5 (±3.5)	3	45
T-52–61	15.5 (±7.7)	5	60
E-52–61	21.8 (4.4)	4	>90
48–62 (DGST-52–61–W)	22.3 (2.9)	3	>90
46–61 (NTDGST-52–61)	22.6 (2.6)	11	>90

Summary of results of all experiments on the persistence of the indicated ¹²⁵I-labelled peptides. Experiments using all three lymphoma cells are included, although most were done in C3.F6 cells, as described in the legends for Figs 1 and 2. With a given peptide, no differences in half-life were found among the three cell lines. Indicated are the mean values, with s.d., and the number of individual experiments (n). '% Stable' is the fraction of labelled peptide bound to I-A^k that ran in the position of a stable dimer on SDS–PAGE. The results of % stability have been published previously¹⁵. For a given peptide–I-A^k complex in APCs, the percentage found in the stable or unstable forms at different times in culture did not vary significantly. The data from three representative experiments are: peptide K-52–61 gave a half-time (t_{1/2}) of persistence of 11.5 h. The percentage stable at 1, 3, 6, 9 and 12 h was 54, 59, 60, 65 and 61%, respectively. Peptide T-52–61 (t_{1/2} 11.5 h in this experiment) showed % stable, at the same times, of 45, 37, 35, 44 and 45%, respectively. Peptide A-52–61 (t_{1/2} of 12 h) showed % stable, at the same times, of 20, 23, 30, 32 and 30%, respectively. This indicates to us that stable and unstable complexes do not represent two independent populations.

FIG. 1 Half-lives of 46-61 and 52-61 bound to I-A^k of TA3 and CH27 cells (a), in C3.F6 in the presence of excess unlabelled peptide (b) and in cell lysates (c). a, Differences in life time of various peptide-MHC class II complexes expressed by B-lymphoma cells in culture. The number of peptide-class II MHC complexes were quantified by immunoprecipitation at different time points from cells cultured in ¹²⁵I-labelled peptides. Results for 46-61 and 52-61 in CH27 cells or TA3 cells in one representative experiment of 10 carried out with the various B lymphomas are shown (see Table 1 for a summary). b, The persistence of ¹²⁵I-labelled 46-61 or ¹²⁵I-52-61 bound to I-A^k of C3.F6 is not changed by culturing in a vast excess of HEL or unlabelled 46-61. C3.F6 were incubated first with labelled peptides, followed by an excess of HEL or unlabelled 46-61. c, The half-life of I-A^k peptide complexes in cell lysates is long. The cells were lysed at 4 h, unlabelled peptides were added and the lysate was incubated; I-A^k-labelled peptides were immunoprecipitated at the indicated times.

METHODS. a, Synthetic peptides were labelled with ¹²⁵I to a specific activity of $2-4 \times 10^9$ c.p.m. μg^{-1} by the chloramine-T method and then purified on C₁₈ reverse-phase HPLC. Cells were exposed to 1,000 rad of gamma irradiation then washed in DMEM containing 2% FCS and resuspended at 5×10^6 cells ml^{-1} . A total of $4-8 \times 10^8$ c.p.m. of labelled peptide was added to each set of 5×10^7 cells, at 37 °C, for 4 h. After washing in medium the cells were resuspended at 1×10^6 cells ml^{-1} in 10-ml flasks. The total uptake of peptide was $\sim 1-5\%$ of the input c.p.m.; the amount precipitable with I-A^k ranged from 1 to 10% of the cell-associated radioactivity. At different times, cells from individual flasks were lysed and complexes immunoprecipitated as previously described¹⁵. Briefly, the cells were collected, washed and then lysed in 0.5 ml 1% Triton X-100, 10 mM iodoacetamide, 1 mM phenylmethylsulphonyl fluoride, and 20 mg ml^{-1} leupeptin. After a 40-min pre-clear with 25 μl 50% Protein-A-Sepharose in PBS containing 1% Triton-X-100, the peptide-I-A^k complexes were immunoprecipitated, using 50 mg ml^{-1} final of monoclonal antibody specific for the I-A^k β -chain. In this case the antibody was 40.F²², although in similar experiments we have used 10.3.6.2 (ref. 23) with essentially identical results. Lys-



ates were incubated with antibody for 2 h with Protein-A-Sepharose added for the last hour of incubation. The complexes were spun, washed and eluted for 1 h at room temperature in 2% SDS loading buffer before analysis by SDS-PAGE. The amount of labelled peptide at each time point was determined by direct gamma counting of excised gel fragments. Half-lives were determined by regression analysis. In b, C3.F6 were incubated with ¹²⁵I-labelled 46-61 or 52-61 for 4 h as described for a. We estimated that at 1 h, 10^7 C3.F6 cells contained 1 fmol of [¹²⁵I]-46-61 or [¹²⁵I]-52-61 bound to I-A^k. The 10^7 cells were then washed and incubated with 700 nmol of HEL and with 1 μmol of 46-61 peptide, expecting to contain 7 nmol of HEL and 1 nmol of 46-61 respectively; this represents a 7 and 1 million-fold excess over the 1 fmol of peptide bound to I-A^k respectively. Aliquots of 10^7 C3.F6 were taken at the indicated times and the I-A^k immunoprecipitated with 10.3.6.2 monoclonal antibody. In c, C3.F6 were incubated with [¹²⁵I]-52-61 or 4 h, as detailed above. After 4 h, the cells (16×10^7) were lysed in 16 ml of lysis buffer. The amounts of [¹²⁵I]-52-61 or [¹²⁵I]-46-61 contained in the lysate were about 30 pmol and 10 pmol respectively. A 10,000-fold excess of nonradioactive peptides was then added to the lysate which was incubated for the times indicated. Aliquots of lysate corresponding to 10^7 cells were collected, immunoprecipitated with 10.3.6.2 antibodies and analysed as detailed in a. Similar experiments at 37 °C resulted in half-dissociation times of 18 and 20 h for 46-61 and 52-61, respectively.

FIG. 2 Culture of C3.F6 with HEL increases the half-life of I-A^k. C3.F6 cells were pulse labelled with [³H]-leucine in the absence or presence of HEL. After a 30-min pulse the cells were cultured in regular media (with or without HEL), and at the indicated times the I-A^k molecules precipitated from aliquots with 10.3.6.2 monoclonal antibody. The amount of immunoprecipitated I-A^k was estimated from eluates on Protein-A-Sepharose beads. The eluates were also analysed by SDS-PAGE. The radioactivity in the stable (that is, $\alpha\beta$ -dimer) and unstable (α -plus β -bands) was measured by counting excised gel fragments, and the percentage in each was calculated. The figure shows the mean results of four different experiments. In each experiment duplicate or triplicate observations were made at each time point. The indicated point represents means of two to four experiments. Variations of the mean ranged over 5-20%. Because the degree of uptake of [³H]-leucine varied among the four experiments, the first time point (3 h) was normalized to 100%. a, Half-life of total [³H]-leucine-labelled I-A^k. b and c, Half-lives of the individual stable and unstable components in cells cultured without or with HEL, respectively. The percentage of stable molecules in C3.F6, not cultured with HEL, was 40, 49, 50, 48 and 51% at 3, 6, 8, 9 and 12 h, respectively. For C3.F6 cultured with HEL the percentage of stable molecules, at the same times, was 66, 70, 78, 76 and 77% respectively. The $t_{1/2}$ values for I-A^k calculated by linear regression analysis were: for C3.F6 cultured in regular media without HEL 12.6 h, with 10.2 h for the unstable and 20.3 h for the stable components. For C3.F6 cultured in media with HEL 23.0 h with 10.9 h for the unstable and 43.0 h for the stable component.

METHODS. C3.F6 cells were cultured at 2×10^7 in leucine-deficient DMEM for 1 h, after which [³H]-leucine (Amersham, 150 Ci mmol^{-1}) was added for 30 min. Aliquots contained 800 μCi per 12×10^7 cells in 6 ml, in T-25 flasks. One aliquot of the cells contained 2 mg ml^{-1} of HEL. After 30 min nonradioactive leucine was added, cells were collected, washed and resuspended in media (without or with HEL, as above), and incubated at 37 °C. At the times indicated, aliquots of 2×10^7 were collected, washed and immunoprecipitated as indicated in Fig. 1.

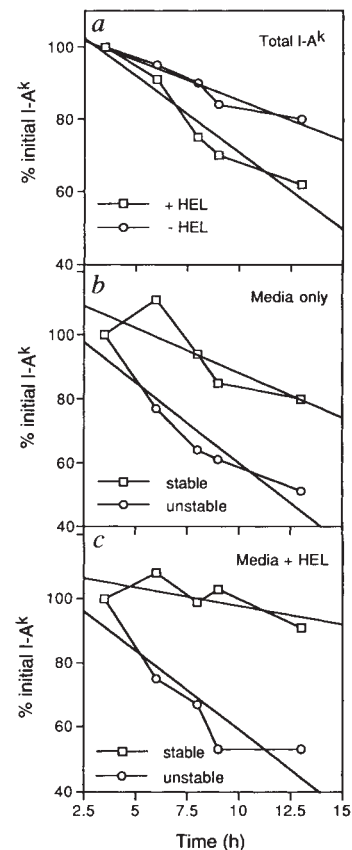


TABLE 2 Loss of immunogenicity of peptide 52–61

Time (h)	IL-2 induced with:	
	46–61 (IL-2 units)	52–61 (IL-2 units)
1	40	18
6	35	1
20	45	0

The immunogenicity of peptide 46–61 persists for long periods whereas that of 52–61 rapidly decays. Irradiated C3.F6 cells were incubated with 46–61 (10×10^6 cells in 5 ml with 1 μ M) or with 52–61 (10×10^6 in 5 ml with 50 μ M) for 4 h at 37 °C. After 4 h the cells were washed, resuspended in medium without peptide and incubated for the indicated time periods. The cells were then fixed in 1% paraformaldehyde (at room temperature for 15 min, then incubated in 0.4 M lysine for 15 min, washed and stored at 4 °C). Cells were assayed with the BD-4 T-cell hybridoma and tested for interleukin-2 (IL-2) production. The figures indicate units of IL-2 assayed by conventional methods on the IL-2 indicator CTLL cells⁹. The BD-4 hybridoma was derived from lymph nodes of CBA mice immunized with 52–61 peptide in complete Freund's adjuvant. Before this experiment identical protocols were evaluated using input amounts of 46–61 of 10 μ M, 1 μ M and 0.3 μ M with identical results; there was no drop in the T-cell response during the first 20 h of incubation. The doses of 10 and 1 μ M of 52–61 showed similar rapid loss of the T-cell response. The doses of 10–1 μ M gave responses in the linear portion of the dose–response curve. In a previous study²⁴ a shorter persistence of peptide was found but under different conditions of incubations (that is, limiting amounts of peptide for only 30 min).

stable molecules (90%), with markedly prolonged half-life. The contribution of the amino acid at position 51 to both the stability and persistence of the complex can be cancelled if the amino-acid side chain is hydrophobic (Table 1 legend, and our unpublished studies), or contains a positive charge (Table 1 legend).

Our results indicate that APCs discriminate between different I–A^k–peptide complexes as evidenced by different times of persistence. These differences inevitably select for long-lived peptides and affect both the potency and immunodominance of the complexes recognized by T cells. The differences in persistence may reflect conformational changes in the I–A^k–peptide complex read out by the propensity of the peptide to form stable SDS–PAGE molecules. Live APCs may sense these differences in structure and handle the complexes in different manners (that is, for example, intracellular turnover rate, recycling, different traffic patterns). For complexes of intermediate stability the correlation to short half-life is not strict. We find a correlation only between peptides having a high propensity to stabilize the I–A^k dimer and the longer half-life. □

Received 18 April; accepted 29 July 1994.

- Rosenthal, A. S., Barcinski, M. A. & Blake, J. T. *Nature* **267**, 156–159 (1977).
- Katz, M. E., Maizels, R. M., Wicker, L., Miller, A. & Sercarz, E. E. *Eur. J. Immun.* **12**, 535–540 (1982).
- Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Unanue, E. R. *Nature* **317**, 359–361 (1985).
- Buus, S., Sette, A., Colon, S. M., Miles, C. & Grey, H. M. *Science* **235**, 1353–1358 (1987).
- Adorini, L., Appella, E., Doria, G. & Nagy, Z. A. J. *exp. Med.* **168**, 2091–2104 (1988).
- Allen, P. M., Strydom, D. J. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2489–2493 (1984).
- Allen, P. M., Matsueda, G. R., Evans, R. J., Dunbar, J. B. jun., Marshall, G. & Unanue, E. R. *Nature* **327**, 713–715 (1987).
- Nelson, C. A., Roof, R. W., McCourt, D. W. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7380–7383 (1992).
- Haughton, G., Arnold, L. W., Bishop, G. A. & Mercolino, T. J. *Immun. Rev.* **93**, 35–51 (1986).
- Glimcher, L. H. et al. *J. Immun.* **130**, 2287–2294 (1983).
- Wade, W. F. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6297–6301 (1989).
- Lanzavecchia, A., Reid, P. A. & Watts, C. *Nature* **357**, 249–252 (1992).
- Pedrazzini, T., Sette, A., Albertson, M. & Grey, H. M. *J. Immun.* **146**, 3496–3501 (1991).
- Adorini, L., Appella, E., Doria, G., Cardinaux, F. & Nagy, Z. A. *Nature* **342**, 800–804 (1989).
- Nelson, C. A., Petzold, S. J. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1227–1231 (1993).
- Dornmair, K., Rothenhausler, B. & McConnell, H. M. *Cold Spring Harb. Symp. quant. Biol.* **54**(1), 409–416 (1989).
- Germain, R. N. & Hendrix, L. R. *Nature* **353**, 134–139 (1991).
- Sadegh-Nasseri, S. & Germain, R. N. *Nature* **353**, 167–170 (1991).
- Germain, R. N. & Rinker, A. G. *Nature* **363**, 725–728 (1993).
- Stern, L. J. & Wiley, D. C. *Cell* **68**, 465–477 (1992).

- Riberdy, J. M., Newcomb, J. R., Surman, M. J., Barbosa, J. A. & Creswell, P. *Nature* **360**, 474–477 (1992).
- Pierres, M., Devaux, C., Dosseto, M. & Marcheto, S. *Immunogenetics* **14**, 481–495 (1981).
- Oi, V. T., Jones, P. P., Goding, J. W. & Herzenberg, L. A. *Curr. Topics Microbiol. Immun.* **81**, 115–129 (1978).
- Harding, C. V., Roof, R. W. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4230–4234 (1989).

ACKNOWLEDGEMENTS. We thank A. Combs and B. Deck for technical assistance, and N. Viner and P. Allen for their comments and advice. Our work is supported by grants from the NIH.

Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia

Francis Rousseau, Jacky Bonaventure, Laurence Legeal-Mallet, Anna Pelet, Jean-Michel Rozet, Pierre Maroteaux, Martine Le Merrer & Arnold Munnich*

Service de Génétique et Unité de Recherches sur les Handicaps Génétiques de l'Enfant, INSERM U. 393, CNRS ER 88, Institut Necker, Hôpital des Enfants Malades, 149 rue de Sèvres, 75743 Paris Cédex 15, France

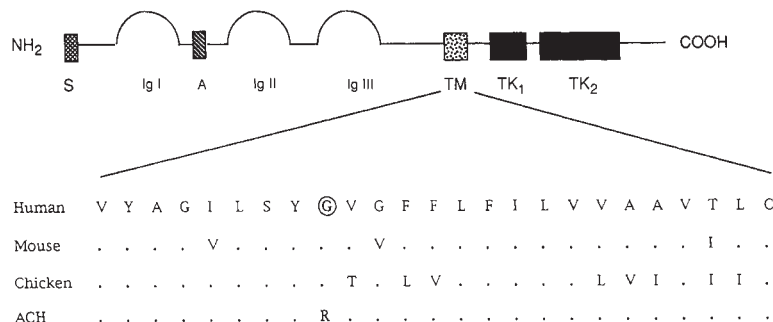
ACHONDROPLASIA, the most common cause of chondrodysplasia in man (1 in 15,000 live births), is a condition of unknown origin characterized by short-limbed dwarfism and macrocephaly^{1,2}. More than 90% of cases are sporadic and there is an increased paternal age at the time of conception of affected individuals, suggesting that *de novo* mutations are of paternal origin. Affected individuals are fertile and achondroplasia is transmitted as a fully penetrant autosomal dominant trait, accounting for rare familial forms of the disease (10%)^{3–6}. In contrast, homozygous achondroplasia is usually lethal in the neonatal period and affects 25% of the offspring of matings between heterozygous achondroplasia parents. The gene responsible for achondroplasia has been mapped to chromosome 4p16.3 (refs 7, 8); the genetic interval encompassing the disease gene contains a member of the fibroblast-growth-factor receptor (FGFR₃) family which is expressed in articular chondrocytes. Here we report the finding of recurrent missense mutations in a CpG doublet of the transmembrane domain of the FGFR₃ protein (glycine substituted with arginine at residue 380, G380R) in 17 sporadic cases and 6 unrelated familial forms of achondroplasia. We show that the mutant genotype segregates with the disease in these families. Thus it appears that recurrent mutations of a single amino acid in the transmembrane domain of the FGFR₃ protein account for all cases (23/23) of achondroplasia in our series.

Individuals with achondroplasia (ACH) have growth cartilage of their long bones that undergoes minimal proliferation⁹, but the cartilage of their remaining epiphyses is normal, implicating a possible abnormality in a growth factor or its chondrocyte receptor⁹.

The gene for ACH has been mapped to chromosome 4p16.3 by linkage in ACH families^{7,8} and the disease locus has been subsequently mapped distal to D4S43 (maximum pairwise lod score at D4S111: $Z_{\max} = 3.01$ at $\theta = 0$ in six ACH families; results not shown). Interestingly, the gene encoding FGFR₃ maps to the same region¹⁰. FGFR₃ is a tyrosine kinase receptor with a large glycosylated extracellular ligand-binding domain containing three immunoglobulin-like repeats, a single hydrophobic transmembrane region and a cytoplasmic domain that contains a tyrosine kinase catalytic domain disrupted by a hydrophobic

* To whom correspondence should be addressed.

FIG. 1 Schematic representation of FGFRs. S, signal sequence; Ig, immunoglobulin-like domains (I, II and III); A, acidic region; TM, transmembrane domain; TK, tyrosine kinase domains 1 and 2. A comparison of the amino-acid sequence of the transmembrane domains of human, mouse and chicken FGFR₃ is also shown. Dots indicate identical residues. Circled G shows the mutated amino acid.



sequence¹¹ (Fig. 1). The mouse homologue of FGFR₃ is expressed in cartilage and mediates the mitogenic effect of basic fibroblast growth factor on chondrocytes and its inhibitory effect on terminal chondrocyte differentiation and cartilage-matrix calcification¹²⁻¹⁴. FGFR₃ was thus regarded, by virtue of both its position and function, as an important candidate gene in ACH.

Preliminary studies indicated that no major deletion or rearrangement of the FGFR₃ locus had occurred in our series (not shown). Our strategy, therefore, was to analyse the coding sequence of FGFR₃ for point mutations or minute changes by using a combination of single-strand conformation polymorphism (SSCP) and direct sequencing analyses of the extracellular, transmembrane and tyrosine kinase domains of the protein. Abnormal patterns of migration indicative of sequence variations in the transmembrane domain were observed in probands from 6/6 unrelated families linked to chromosome 4p16.3 and in 17/17 sporadic ACH cases. The pattern of migration was identical in 22/23 probands (Fig. 2a, c, d) and direct sequence analysis showed heterozygosity for the same missense mutation (G→A), which converted a glycine into an arginine in the protein (G380R). The other abnormal SSCP pattern (Fig. 2b) was due to a transversion at the same codon (G→C), which resulted in the same G380R substitution. The base substitutions, which involved a CpG doublet, were absent in 120 normal controls ($P < 0.001$).

Additional evidence that mutation of the FGFR₃ gene is significant in ACH was provided by (1) the cosegregation of the base substitution with the disease in each of six families (Fig. 2a); (2) its *de novo* occurrence in sporadic cases (Fig. 2c, d); and (3) the observation of homozygosity for the G380R mutation in a case of lethal homozygous ACH (not shown). But no mutation of the transmembrane domain was found in a mild variant of ACH (hypochondroplasia; results not shown), so other domains

of the FGFR₃ gene or a different locus may be involved in hypochondroplasia.

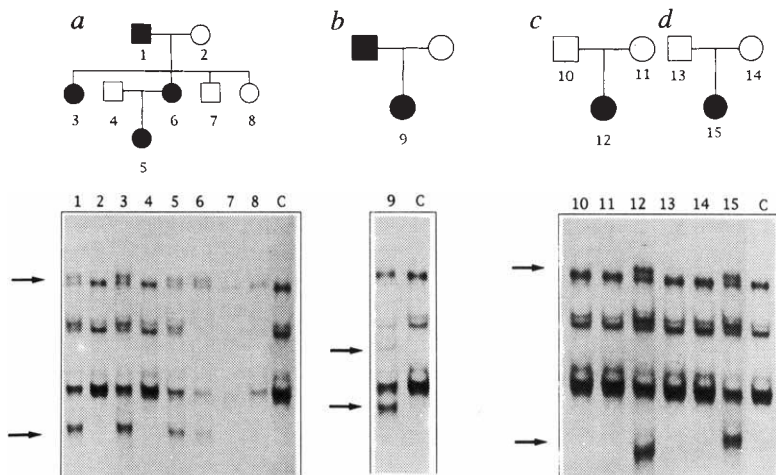
Deletions of chromosome 4p encompassing the FGFR₃ gene cause Wolf-Hirschhorn syndrome, a condition whose clinical expression (growth failure, mental retardation, cardiac and bone malformations) is very different from ACH. The striking differences between the two diseases suggest that ACH does not result from a gene dosage effect, but rather from the dominant negative effect of the mutant gene product on the wild-type counterpart. In keeping with this, disruption of receptor dimerization is an attractive model for a dominant negative mutation as ligand binding is known to induce FGFR₃ dimerization. Hence a mutation in the transmembrane domain may hinder the conformational change required for FGFR₃ dimerization with itself or with other FGFRs. Alternatively, disruption of the α -helical structure of the transmembrane domain may prevent the activation of tyrosine kinase function, especially as the G380R mutation involves a highly conserved amino acid^{15,16} (Fig. 1) and introduces a charged residue into a highly hydrophobic domain of the receptor.

Why should 23/23 unrelated probands bear the same mutation? Perhaps the CpG dinucleotide at codon 380 promotes independent mutations in a residue critical for signal transduction and/or receptor anchorage (especially if the cytosine residue of the CpG doublet is methylated), whereas other mutations involving the transmembrane domain are less damaging. In fact, mutations in other tyrosine kinase receptors, for example in the *RET* proto-oncogene or in platelet-derived growth factor, have a neutral, activating or inhibitory effect, depending on their position in the transmembrane domain^{17,18}.

We conclude that FGFR₃ is the gene responsible for ACH, having shown that a mutation in a transmembrane domain of this fibroblast growth factor receptor results in a skeletal growth defect. Remarkably, independently recurring mutations at a

FIG. 2 Single-strand conformation polymorphism (SSCP) analysis of the transmembrane domain of the FGFR₃ gene in four ACH probands (a-d). Numbering in top panels corresponds to lane numbering in bottom panels. Arrows indicate abnormal bands in probands; C, controls.

METHODS. Oligonucleotide primers for the polymerase chain reaction (PCR) amplification of a fragment (159 bp) containing the entire transmembrane domain were devised according to ref. 12. Forward primer: 5'-TGAC-GAGGCGGGCAGTG-3'; backward primer: 5'-AGCGGGGAA-GCGGGAGAT-3'. Genomic DNA (100 ng) was amplified by PCR (30 cycles at 94 °C for 40 s, 50 °C for 20 s, 72 °C for 20 s) in 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 25 pmol primers, 5% DMSO, 200 μ M of each deoxynucleotide triphosphate, 1 U *Taq* polymerase, using [α -³²P]dCTP in 25 μ l. PCR products were heated for 10 min at 95 °C, loaded onto a Hydrolink MDE gel (Bioprobe systems) and electrophoresed for 13 h at 4 W. Gels were dried and autoradiographed.



single codon were responsible for all the cases of ACH in our series. As no animal model for ACH is available at present, the production of transgenic animals bearing heterozygous mutations in the transmembrane domain of the FGFR₃ gene should open up the as-yet unexplored field of cell biology and molecular pharmacology of chondrodysplasia and lead to the development of effective treatment of this frequent cause of dwarfism. □

Received 21 July; accepted 18 August 1994.

1. Maroteaux, P. & Lamy, M. *Clin. Orthop.* **33**, 91–103 (1964).
2. Oberklaid, F., Danks, D. M., Jensen, F., Stace, I. & Rosshandler, S. *J. med. Genet.* **16**, 140–146 (1979).
3. Murdoch, J. L. et al. *Ann. hum. Genet.* **33**, 227–244 (1970).
4. Orioli, I. M., Castilla, E. E. & Barbosa-Neto, J. G. *J. med. Genet.* **23**, 328–332 (1986).
5. Stoll, C., Dott, B., Roth, M. P. & Alembik, Y. *Clin. Genet.* **35**, 88–92 (1989).

6. Gardner, R. J. M. *Clin. Genet.* **11**, 31–38 (1977).
7. Le Merrer, M. et al. *Nature Genet.* **6**, 318–321 (1994).
8. Velinov, M. et al. *Nature Genet.* **6**, 314–317 (1994).
9. Stanesco, R., Stanesco, V. & Maroteaux, P. *Am. J. med. Genet.* **37**, 412–421 (1990).
10. Thompson, L. et al. *Genomics* **11**, 1133–1142 (1991).
11. Ullrich, A. & Schlessinger, J. *Cell* **61**, 203–212 (1990).
12. Keegan, K., Johnson, D., Williams, L. T. & Hayman, M. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1095–1099 (1991).
13. Kato, Y. & Iwamoto, M. *J. biol. Chem.* **265**, 5903–5909 (1990).
14. Iwamoto, M., Shimazu, A., Nakashima, K., Suzuki, F. & Kato, Y. *J. biol. Chem.* **266**, 461–467 (1991).
15. Ornitz, D. M. & Leder, P. *J. biol. Chem.* **267**, 16305–16311 (1992).
16. Pasquale, E. B. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5812–5816 (1990).
17. Williams, L. T. *Science* **243**, 1564–1570 (1989).
18. Edery, P. et al. *Nature* **367**, 378–380 (1994).

ACKNOWLEDGEMENTS. We thank J. Frézal and M. Hayden for discussion, C. Gassel and S. Strautnieks for help in preparing the manuscript, and C. Stoll, M. Sanak and the Association des Personnes de Petite Taille for their cooperation. This work was supported by the Caisse Nationale d'Assurance Maladie des Professions Indépendantes, the Association Française contre les Myopathies and the Groupement de Recherches et d'Etudes sur les Génomes.

A cyclin associated with the CDK-activating kinase MO15

Toml P. Mäkelä, Jean-Pierre Tassan*,
Erlch A. Nigg*, Séverine Frutiger†,
Graham J. Hughes† & Robert A. Weinberg‡

Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA

* Swiss Institute for Experimental Cancer Research (ISREC), 155 Chemin des Boveresses, CH-1066 Epalinges, Switzerland

† Centre Médicale Universitaire, Université de Genève, Département de Biochimie Médicale, 1 rue Michel-Servet, CH-1211 Genève 4, Switzerland

THE eukaryotic cell cycle is regulated by the sequential activation of cyclin-dependent kinases (CDKs)¹. CDK activation is dependent on cyclin binding^{2–4} and phosphorylation of a conserved threonine (T161 in Cdc2)^{5–9} mediated by the CDK-activating kinase CAK⁹. A CDK-related kinase, MO15 (ref. 10), has been identified as the catalytic subunit of CAK (refs 11–13). Here we use a yeast two-hybrid screen to show that a new human cyclin (cyclin H) is a MO15-associated protein. Cyclin H is a major MO15 partner *in vivo* and enhances the kinase activity of MO15 towards Cdk2/cyclin A. These findings demonstrate that a cyclin/kinase complex can function as a regulator of other cyclin/kinase complexes, and suggest that cyclin/kinase cascades may exist.

A yeast two-hybrid screen¹⁴ was used to identify proteins interacting with the human homologue of *Xenopus* MO15 (ref. 10). The assay was based on reconstituting a functional transcriptional activator from two separate fusion proteins, using in this case a DNA-binding LexA–MO15 as 'bait', and a transactivation domain fused with HeLa cell complementary DNAs as 'prey'¹⁵. One class of positive clones was found to contain an open reading frame encoding a sequence of 323 amino acids having significant sequence homology to several cyclins (Fig. 1a). The strongest similarities were found with *Schizosaccharomyces pombe* mcs2 (ref. 16; 31% identity) and *Saccharomyces cerevisiae* Ccl1 (ref. 17; 31% identity) cyclins, and with mammalian cyclin C (ref. 18; 23% identity; Fig. 1b). On the basis of these homologies we have named this protein cyclin H.

The specificity of the cyclin H/MO15 association was supported by the inability of the cyclin H prey to interact with LexA–bicoid or LexA–MO15Δ1–132 (Fig. 2). Further insight into the cyclin H/MO15 interaction was provided by MO15

bait with specific mutations. The kinase-inactive MO15–K41M retained the interaction, which thus is not dependent on kinase activity. The T170E mutation targets a potential phosphorylation site in MO15 conserved in CDKs; this mutation abolished the interaction. Interestingly, a corresponding T160E change in Cdk2 does not abrogate cyclin binding¹⁹.

To test whether MO15 and cyclin H associated with one another in mammalian cells, Myc-tagged MO15 (MO15-3M) and HA-tagged cyclin H (CycH-HA) proteins were expressed in COS cells (Fig. 3). Anti-Myc immunoprecipitates from cells transfected with MO15-3M plasmid yielded a strong doublet band corresponding to a relative molecular mass of 41K. In cells cotransfected with MO15-3M and CycH-HA plasmids, we observed an additional strong band at 39K (as well as a weak band at 34K), which comigrated with a 39K band observed in anti-HA immunoprecipitates from cells transfected only with CycH-HA plasmid, and thus represents MO15-associated cyclin H. Similarly, we could also co-immunoprecipitate the MO15–3M protein with cyclin H. However, only the slower migrating of the two MO15 bands co-precipitated with cyclin H (Fig. 3a). We extended these experiments by demonstrating that untagged cyclin H and MO15 could each be co-precipitated with their respective tagged partners (Fig. 3b). We concluded that native human MO15 and cyclin H proteins have electrophoretic mobilities of 37K and 34K, respectively.

Taken together, these results indicate that MO15 and cyclin H associate efficiently in mammalian cells. The specificity of this interaction is supported by the inability of cyclin H to interact with Cdc2, Cdk2 or MAPK, and of MO15 to interact with cyclin A, E or C in similar conditions (data not shown). In addition, our results suggest that the slower migrating band of the 39K MO15 doublet may represent a phosphorylated and possibly activated form of MO15 favoured as the cyclin H partner. Indeed, in cotransfections a larger proportion of MO15 was found in this slower migrating form (Fig. 3a, b).

To extend our studies to endogenous CAK, we immunoprecipitated MO15 from labelled HeLa cells using anti-MO15 antibodies¹⁰ (Fig. 3c, lane 1). As reported²⁰, two polypeptides of 34K and 32K co-immunoprecipitated in near-stoichiometric amounts with endogenous MO15. We microsequenced one of these (p34), which had a size similar to that of cyclin H; the MO15 band from the same experiment was used as a control (Fig. 3c). All three sequenced peptides of p34 aligned with the deduced cyclin H sequence (Fig. 3c). This demonstrated the identity of p34 and cyclin H and indicated that MO15 and cyclin H are found in physical complexes in physiological conditions.

The *Xenopus* MO15 expressed as a bacterial glutathione S-transferase(GST)-fusion protein lacks kinase activity toward its substrate, Cdk2 (ref. 12). However, preincubation of GST–

‡ To whom correspondence should be addressed.

a

```

GGACGCTGATGCGTTTGGGTTCTCGTCTGCAGACCCCTCTGGACCTGGTCACGATGCCATA 60
M Y H N S S Q K R H W T F S S E E Q L A 20
ATGTACCACACAGTAGTCAGAACGCCACTGGACCTTCTCCAGCGAGGAGCAGCTGGCA 120
R L R A D A N R K F R C K A V A N G K V 40
AGACTGGCGGCTGACGCCAACCGCAATTCAGATGCAAGCCGTGGCCACGGGAAGTT 180
L P N D P V F L E P H E E M T L C K Y Y 60
CTTCGGAATGATCCAGCTTCTTCTGAGCCTCATGAAGAAATGACACTTGCATAACTAT 240
E K R L L E F C S V F K P A M P R S V V 80
GAGAAAGGTTATTTGGAATTCGTGCTGGTGTAAAGCCAGCAATGCCAAGATCTGTGTG 300
G T A C M Y F K R F Y L N N S V M E Y H 100
GGTAGCGCTTGTATGATTTCAAGACCTTTTCTTAACTACAGTAATGGAATATCAC 360
P R I I M L T C A F L A C K V D E F N V 120
CCCAGGATAATATGCTACTTGTGCAATTTTGGCCGCAAGATAGATGAATCAATGTA 420
S S P Q F V G N L R E S P L G Q E K A L 140
TCTAGCTCCAGTTTGTGGAAACCTCCGGAGAGTCTCTGGACAGGAGAGGCACTT 480
E Q I L E Y E L L L I O Q L N F H L I V 160
GAACAGATCTGGAATATGAATCTTCTTATACAGCACTTAATTTCCACCTTATGTG 540
H N P Y R P F E G F L I D L K T R Y P I 180
CACAACTCTTACAGACCATTTAGGCGTCTCATGACTTAAAGACCGCTATATCCATA 600
L E N P E I L R K T A D D F L N R I A L 200
TTGGAGAAATCCAGAGATTTTGGAGAAACAGCTGATGACTTCTTAAATAGAATGCAATG 660
T D A Y L L Y T P S O I A L T A I L S 220
ACCGAGCTTACCTTTTATACACACTTCCCAAGTGGCCTGACTGCCATTTATCTAGT 720
A S R A G I T M E S Y L S E S L M L K E 240
GCCCTCCAGGCTGGAATTTACTATGGAAGTTATTTATCAGAGAGTCTGATGCTGAAAGAG 780
N R T C L S Q L L D I M K S M R N L V K 260
AACAGAACTTGCCTGTACAGTTTACTATGAAAGCATGAGAACTTGTAGTAAG 840
K Y E P P R S E E V A V L K Q K L E R C 280
AAGTATGAACCCAGATCTGAAGAAGTGTCTGTTCTGAAACAGAAGTTGAGAGGATGT 900
H S A E L A L N V I T K K R K G Y E D D 300
CATCTGCTGAGCTTGCACTTAACCTAATCAGAGAAAGGAGGATGATGAAGATGAT 960
D Y V S K K S K H E E E E W T D D D L V 320
GATTACGCTCTCAAGAAATCCAAACATGAGGAGGAAGAATGGACTGATGACGACCTGGTA 1020
E S L * 323
GAATCTCTTAACCATTTGAAGTTGATTCTCAATGCTAATCAAGAGAAGTAGGAA 1080
GCATATCAACGTTTAACTTTATTTAAAGTATAATGTGAAACATAAAATATATATAA 1140
ACTTTCTATTGTTTCTTTCCCTTTCACAGTAACCTTTATGTAAATAAACCATCTTCAA 1200
AAGAAAAAAGAAAAA 1219

```

FIG. 1 Sequence of cyclin H. **a**, Nucleotide and deduced amino-acid sequence of cyclin H. Arrow indicates the start of cDNA clone F11 isolated in the two-hybrid screen. Further 5' cDNA sequence was

b

```

mcs2 MSADKFRDSTHYRDWIF..TEEDLSKTRAKVNEKFTNIVRERMLELSLQNKESLEVLPP
cycH MYHNSQKRHWTF..SSEEQARLRADANKF...RCKAVANGKVLNPDVVF....
cycC MVAPRPLRRVLFYQGLKLSMAGN...F...WQSSHYLOWLDKQDLL....

mcs2 TLTVEEELVLYNYSFQLNA..LSSAL.....SLPHTIRSTAILFFKRFYILNSVMEY
cycH .LEPHEMTLCKYYEKRLLE..FCSVFKP....AMP RSVVGTACHYFKRFYILNSVMEY
cycC .KERQDKLFLSEEEYWLQIFFTNVIQALGHEHLKQQVIATATVYKRFYARYSLKSI

mcs2 SPKIIISFTSLFLATKCNDRHISIEQFCCKNPKT.....T.....PEEVLEYEFN
cycH HPRIIMLTCAFLACKVDEFNVSPQFVGNLRESPLGQEK.....LEQILEYELL
cycC DPVLMAPTCLVFLASKVEFGVSNTRLIAAATSVLKTRFSYAFKPEFFYRMHILECEFY

mcs2 VCQSLKWDLYVWLPFRPLDGLLDCQTVLPKVAVEKFYECHDLSKKFLIETLHSDIYFLH
cycH LIQQLNFHLIVNHPYRPFEGFLIDLKTRYP ILENPEIL..RKATDDFLNRIALTOAYLLY
cycC LLELMDCCLIVYHPYRPLQYVQD.....MQGEDML..LPLAWRIVNDTYRTOLCLLY

mcs2 SPSIIALGAI..Y..HTNPTICLO..YIEAKKIPLEQPLIISANLKAATKFKIEKKKAQD
cycH TPQIALTAI..LSSASRAGITMESYLSLSEMLKENRTCSQLDLMKSMRNLVKYEPFR
cycC PPFMIACLHVACVQKQDARQWFAELSDVMEKILEIRVLKLVQEYQKNFDER....

mcs2 YGRKLYFCMPLRNKSSALYL....KRKAEESTNNKWK...KFSTSSNVLDKNPFE
cycH SEEVAVLKQKLERCHSAELALNVI..TKRKGYYDDVYSKSKSHEEEWTDODLVESL
cycC .KEMATILSKMPKPKPP.....PNSEGEQNGSQNSYSQS

```

obtained from cDNA clone F11-1 isolated from a human fetal liver cDNA library (Clontech). Asterisks indicate stop codons upstream and at the end of the long open reading frame. The underlined sequence indicates the polyadenylation signal. **b**, Three-way comparison of amino-acid sequences of *S. pombe* mcs2 (ref. 16), human cyclin H, and human cyclin C (ref. 18) was compiled from two-way comparisons made with the GAP program of the Genetics Computer Group GCG package²¹. The boxed area indicates the region of high conservation among all cyclins.

MO15 with a cellular extract activates the protein, allowing it to phosphorylate a bacterially produced GST-Cdk2 exclusively on Thr 160; this occurs in either the presence or absence of cyclin A¹². These results suggested that post-translational modification or association with other cellular components was necessary for MO15 activation. We speculated that our cyclin H protein might serve as the cellular component that associates with and functionally activates MO15.

To this end we measured the kinase activities of MO15 or MO15/cyclin H immunocomplexes towards GST-Cdk2 or GST-Cdk2/cyclin A (Fig. 4a). As negative controls we used cyclin H alone or MO15-K41M/cyclin H immunocomplexes. As expected, in cyclin H immunoprecipitates GST-Cdk2 phosphorylation was only observed when MO15 was included (Fig. 4a, α -HA panel). Accordingly, in MO15 immunoprecipitates

kinase activity toward GST-Cdk2/cyclin A was strongest when cyclin H was included (Fig. 4a, α -Myc panel). In parallel with the increase in the phosphorylation of GST-Cdk2, histone H1 kinase activity was increased (Fig. 4a, bottom panel, and Fig. 4b). We concluded that cyclin H enhances the kinase activity of MO15 toward Cdk2/cyclin A, leading to activation of Cdk2 in the presence of an appropriate cyclin partner, in this case cyclin A. However, the modest threefold activation of MO15 by cyclin H may indicate involvement of additional components (such as the MO15-associated p32; ref. 20) in the CAK enzyme *in vivo*.

The structural similarities between MO15 and other CDKs suggested that these kinases might all function in a similar way and have common regulatory mechanisms. Our results support this model, indicating that the cyclin H/MO15 complex functions like other cyclin/kinase complexes. However, MO15

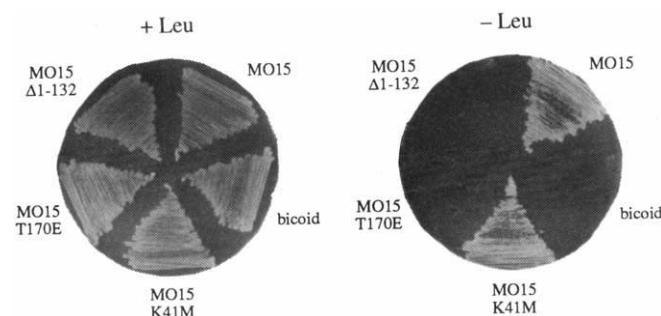
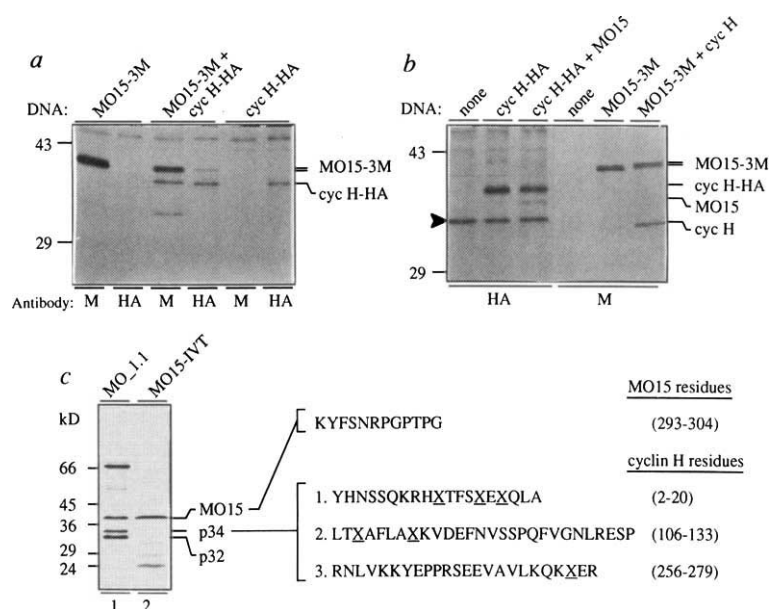


FIG. 2 Interaction of cyclin H with MO15 in yeast. A cyclin H 'prey' plasmid was transformed into yeast cells expressing the indicated proteins as 'bait'¹⁵. Interaction of the DNA-binding bait fusion protein and the transactivating prey fusion protein results in the transcription of the *LEU2* gene and thus enables yeast cells to grow on medium lacking leucine¹⁵. Transformants were restreaked on galactose-containing plates lacking histidine and tryptophan (+Leu) or lacking histidine, tryptophan and leucine (-Leu).

METHODS. The MO15 bait plasmid contains a full-length cDNA of the human homologue of the *Xenopus* MO15²² cloned by PCR from a WI-38 human fibroblast cDNA library (gift from C. Sardet) and inserted into pEG202¹⁵. MO15-K41M and MO15-T170E were generated similarly. The N-terminal deletion MO15Δ1-132 was generated by a *Bam*HI deletion of the MO15 bait plasmid, and the bicoid bait plasmid¹⁵ used as a nonspecific interactor control was provided by R. Brent.

FIG. 3 Association of tagged MO15 and cyclin H in mammalian cells. **a**, COS cells were transfected with plasmids expressing MO15 with three copies of the Myc epitope tag at the C terminus (MO15-3M), or expressing cyclin H with one copy of the HA epitope tag at the N terminus (cycH-HA) as indicated on top. [35 S]methionine-labelled lysates were immunoprecipitated with Myc (lanes M) or HA (lanes HA) antibodies as indicated. **b**, Native cyclin H and MO15 can be detected by association with their respective tagged partners. In addition to plasmids encoding the tagged proteins, plasmids encoding untagged MO15 (MO15) or cyclin H (cycH) proteins were used for transfections as indicated on top. Positions of the 37K native MO15 and 34K cyclin H proteins are indicated on the left; arrowhead marks a nonspecific background band seen in some anti-HA immunoprecipitates. COS transfections and immunoprecipitations have been described²³. **c**, Peptide sequence analysis of MO15 and associated p34. In lane 1 MO15 was immunoprecipitated from [35 S]methionine-labelled HeLa cells using MO1.1 monoclonal antibodies as described²⁰, showing MO15 (comigrating with *in vitro* translated MO15 in lane 2) and associated polypeptides (p32 and p34). Peptide sequences of MO15 and p34 were obtained from a similar large-scale experiment. Underlined Xs in sequences denote uncertainties; residue numbers of MO15 or cyclin H matching the peptide sequences are indicated on the right.

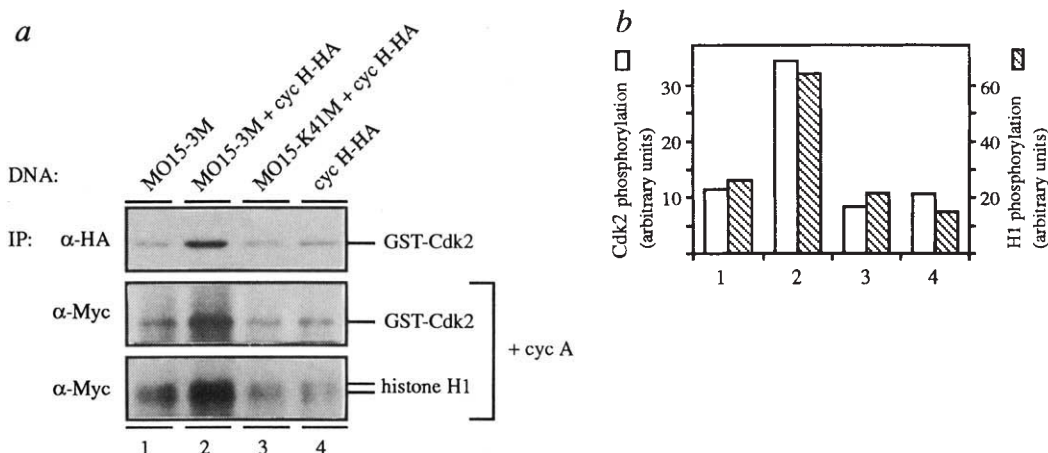
METHODS. For sequence analysis of MO15 and p34, seventy 15-cm² dishes of HeLa cells were lysed and immunoprecipitated with protein G-Sepharose crosslinked²⁴ MO1.1. Then 100–200 pmol proteins were separated on SDS-PAGE, visualized by negative staining²⁵, and cleaved with cyanogen bromide. Resulting peptides were separated



on Tris/tricine-SDS gels²⁶, blotted onto PVDF membranes, visualized with amido black, cut, and subjected to N-terminal sequencing with a model 473A or 477A microsequencer (Applied Biosystems).

FIG. 4 Kinase activity of MO15 and MO15/cyclin H immunocomplexes. **a**, COS cells transfected with plasmids encoding Myc-tagged MO15 (MO15-3M) or HA-tagged cyclin H (cycH-HA) as indicated were used for immunoprecipitation (IP) with either HA (α -HA) or Myc (α -Myc) antibodies. Precipitates were then measured for kinase activity toward GST-Cdk2 or a GST-cdk2/cyclin A complex as indicated. **b**, Densitometric quantification of the relative phosphorylation of GST-Cdk2 (white bars) and histone H1 (striped bars) in the anti-Myc immunocomplexes indicated in **a**.

METHODS. COS cell transfections and immunoprecipitations were done as for Fig. 3, except that the final two washes were in kinase buffer (20 mM Tris, pH 7.5, 5 mM MgCl₂, 2.5 mM MnCl₂, 1 mM dithiothreitol). The resulting immunocomplexes were incubated in kinase buffer supplemented with 50 μ M ATP, 2.5 μ Ci [γ -³²P]ATP, and either 2 μ g bacterially produced GST-Cdk2 purified as described¹² or 2 μ g of a GST-Cdk2/cyclin A complex. This complex was prepared by incubating GST-Cdk2 with cyclin A baculovirus-infected insect cell lysates made in ELB (150 mM NaCl, 50 mM HEPES, pH 7.5, 5 mM EDTA, 0.1% NP-



and its partner cyclin H are unlike the other CDK/cyclin pairs in one important respect; the favoured substrate of MO15/cyclin H is a central component of the core cell cycle machinery itself. This suggests the possible existence of kinase cascades in which the intermediates are CDKs.

We note that cyclin/kinase complexes analogous to cyclin H/MO15 may operate in protozoa. Thus, Ccl1 in budding yeast and mcs2 in fission yeast are the closest homologues of cyclin H. The kinase partner for Ccl1 is Kin28 (ref. 17), and csk1 may be a kinase partner for mcs2 (ref. 16). Both Kin28 and csk1 kinases are most closely related to MO15. Finally, the closest

relative of cyclin H among the human G1 cyclins is cyclin C (ref. 18), for which no kinase partner or activity has so far been described.

Received 5 July; accepted 23 August 1994.

- Norbury, C. & Nurse, P. A. *Rev. Biochem.* **61**, 441–470 (1992).
- Draetta, G. et al. *Cell* **56**, 829–838 (1989).
- Pines, J. & Hunter, T. *Cell* **58**, 833–846 (1989).
- Solomon, M. J., Glotzer, M., Lee, T. H., Philippe, M. & Kirschner, M. W. *Cell* **63**, 1013–1024 (1990).
- Gu, Y., Rosenblatt, J. & Morgan, D. O. *EMBO J.* **11**, 3995–4005 (1992).
- Desai, D., Gu, Y. & Morgan, D. O. *Molec. Biol. Cell* **3**, 571–582 (1992).

5. Gu, Y., Rosenblatt, J. & Morgan, D. O. *EMBO J.* **11**, 3995–4005 (1992).
6. Desai, D., Gu, Y. & Morgan, D. O. *Molec. Biol. Cell* **3**, 571–582 (1992).
7. Ducommun, B. et al. *EMBO J.* **10**, 3311–3319 (1991).
8. Gould, K. L., Moreno, S., Owen, D. J., Sazer, S. & Nurse, P. *EMBO J.* **10**, 3297–3309 (1991).
9. Solomon, M. J., Lee, T. & Kirschner, M. W. *Molec. Biol. Cell* **3**, 13–27 (1992).
10. Shuttleworth, J., Godfrey, R. & Colman, A. *EMBO J.* **9**, 3233–3240 (1990).
11. Fesquet, D. et al. *EMBO J.* **12**, 3111–3121 (1993).
12. Poon, R. Y., Yamashita, K., Adamczewski, J. P., Hunt, T. & Shuttleworth, J. *EMBO J.* **12**, 3123–3132 (1993).
13. Solomon, M. J., Harper, J. W. & Shuttleworth, J. *EMBO J.* **12**, 3133–3142 (1993).
14. Fields, S. & Song, O. *Nature* **340**, 245–246 (1989).
15. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. *Cell* **75**, 791–803 (1993).
16. Molz, L. & Beach, D. *EMBO J.* **12**, 1723–1732 (1993).
17. Valay, J. G., Simon, M. & Faye, G. *J. molec. Biol.* **234**, 307–310 (1993).
18. Lew, D. J., Dulic, V. & Reed, S. I. *Cell* **66**, 1197–1206 (1991).
19. Connell-Crowley, L., Solomon, M. J., Wei, N. & Harper, J. W. *Molec. Biol. Cell* **4**, 79–92 (1993).
20. Tassan, J.-P., Schultz, S., Bartek, J. & Nigg, E. A. *J. cell. Biol.* (in the press).
21. Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387–395 (1984).
22. Levedakou, E. N. et al. *Oncogene* **9**, 1977–1988 (1994).
23. Hinds, P. W. et al. *Cell* **70**, 993–1006 (1992).
24. Harlow, E. & Lane, D. *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1988).
25. Ortiz, M. L., Calero, M., Patron, C. F., Castellanos, L. & Mendez, E. *FEBS Lett.* **296**, 300–304 (1992).
26. Schagger, H. & von Jagow, G. *Analyt. Biochem.* **166**, 368–379 (1987).

ACKNOWLEDGEMENTS. We thank R. Finley and R. Brent for interaction trap reagents; the Harlow laboratory for the GST-cdk2 plasmid; D. Morgan for cyclin A baculovirus; and members of the Weinberg laboratory, especially C. Sardet and R. Medema, for reagents, help and discussion. This work was supported by a grant from the American Cancer Society to R.A.W., T.P.M. is a recipient of a Human Frontiers Science Program Organisation long-term fellowship; J.-P.T. holds a postdoctoral fellowship from the French Cancer League; S.F. and G.J.H. are supported by the Swiss NSF and E.A.N. by the Swiss NSF and the Swiss Cancer League. R.A.W. is an American Cancer Society Research Professor. The GenBank accession number for the cyclin H cDNA sequence is U11791.

p15^{INK4B} is a potential effector of TGF- β -induced cell cycle arrest

Gregory J. Hannon & David Beach*

Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

TRANSFORMING growth factor-beta (TGF- β) inhibits cell proliferation by inducing a G1-phase cell cycle arrest¹. Normal progression through G1 is promoted by the activity of the cyclin-dependent protein kinases CDK4 and CDK6 (ref. 2), which are inhibited by the protein p16^{INK4}. We have isolated a new member of the p16^{INK4} family, p15^{INK4B}. p15 expression is induced ~30-fold in human keratinocytes by treatment with TGF- β , suggesting that p15 may act as an effector of TGF- β -mediated cell cycle arrest. The gene encoding p15 is located on chromosome 9 adjacent to the p16 gene at a frequent site of chromosomal abnormality in human tumours (9p21).

TGF- β is a multifunctional polypeptide that elicits different responses in different cell types and is the prototype of a family of growth factors, differentiation factors and morphogens¹. For most cell types (such as epithelial, endothelial, myeloid and lymphoid) TGF- β is a negative growth factor. Exposure causes cells to cease proliferation and undergo cell cycle arrest in G1 phase¹. Current models suggest that during G1, cyclin-dependent kinases (particularly the cyclin D-associated kinases-CDK4 and CDK6) phosphorylate the product of the retinoblastoma-susceptibility gene, *Rb*, and thus release cells from its growth inhibitory effects^{2–4}. TGF- β treatment causes accumulation of *Rb* in the underphosphorylated state, and expression of *Rb*-inactivating viral oncoproteins prevents TGF- β -induced cell cycle arrest^{5,6}. These results suggested that TGF- β might function by suppressing *Rb* phosphorylation and pointed to the possibility that one result of TGF- β treatment might be inhibition of cyclin-dependent kinases.

To investigate the mechanism by which TGF- β inhibits cell proliferation, we examined α -CDK immunoprecipitates from human keratinocytes which had been arrested in G1 by exposure to TGF- β ^{7,8}. Immunoprecipitates of two G1-specific cyclin kinases, CDK4 and CDK6, contained several associated proteins of low molecular weight (Fig. 1c). These included p16^{INK4}, a previously characterized CDK4 inhibitor⁹, and two additional proteins of relative molecular mass (M_r) ~15,000 and ~15,500. These proteins were not recovered in parallel CDC2 or CDK2 immunoprecipitates (not shown) but were recovered (albeit to a low extent) in α -p16 immunoprecipitates (Fig. 1c), suggesting that p15, p15.5 and p16 might be related. This was confirmed by western blotting of CDK4 and CDK6 immunoprecipitates which demonstrated that p15 and p15.5 were weakly cross reactive with the p16 antiserum (not shown).

To isolate clones encoding putative p16 relatives, we constructed a cDNA library from TGF- β arrested HaCaT cells and probed this library at low stringency with the p16 coding sequence. One clone obtained in this screen encoded a protein of 137 amino acids (predicted M_r 14,700 (14.7K) with homology to p16^{INK4}. On the basis of this homology and the biochemical properties described below, we have named this protein p15^{INK4B} (inhibitor of CDK4-B). The first 50 amino-acid residues of p15 and p16 share ~44% identity (Fig. 1b). This is followed by an 81 amino acid region of ~97% identity after which p15 and p16 diverge. The sequence of p15^{INK4B} can be divided into four ankyrin repeats suggesting that this structural motif is conserved in the INK4 family⁹. *In vitro* translation of the p15^{INK4B} cDNA produced a protein which precisely comigrated with the p15 present in CDK4, CDK6 and p16 immunoprecipitates from TGF- β -arrested HaCaT cells (not shown). Identity of these proteins was confirmed by protease and chemical cleavage mapping (Fig. 1d).

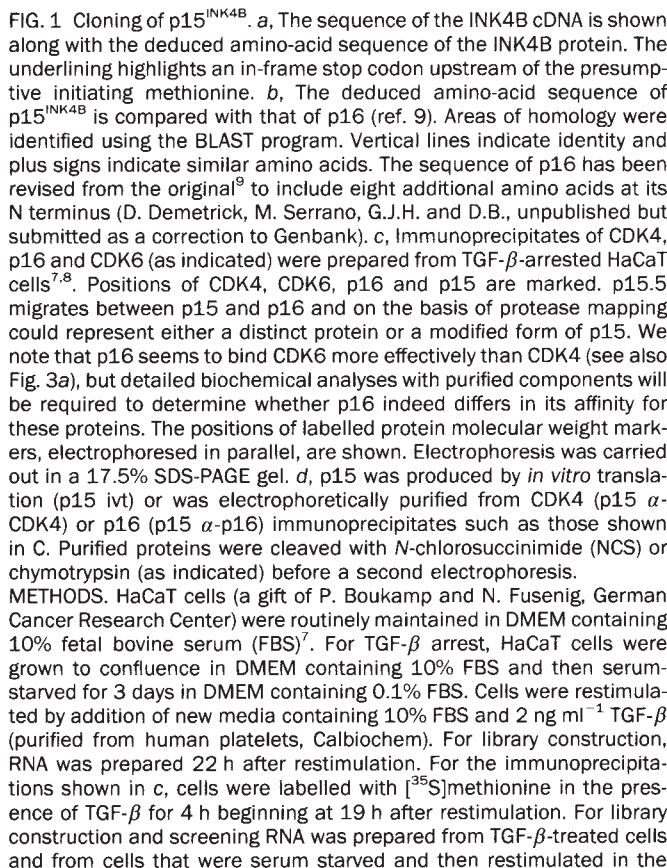
To investigate the functional similarity of p15 and p16, we expressed p15 as a fusion protein in bacteria and tested its ability to bind and inhibit cyclin-dependent kinases. p15^{INK4B} specifically bound CDK4 and CDK6 but did not appreciably bind CDC2, CDK2 or CDK5 (Fig. 2a). To assess the consequences of binding, p15 was added to active cyclin-CDK complexes expressed in insect cells. p15 specifically inhibited the cyclin D-CDK4 and cyclin D-CDK6 enzymes (Fig. 2b, c) but had no effect on CDK2-cyclin A kinase (not shown). Thus, p15^{INK4B} is a functional member of the p16^{INK4} family.

The carboxy-terminal portion of the p15 sequence was reminiscent of the amino-acid sequence predicted from a genomic locus, *MTS2*, which had been previously identified¹⁰. Comparison revealed that the reported *MTS2* sequence encodes the C-terminal 86 amino acids of p15^{INK4B}. *MTS2* lies adjacent to the p16 gene at 9p21, and localization of the p15 gene to this position was confirmed by FISH mapping (D. Demetrick, G.J.H. and D.B., unpublished). Although a fragment of the p15 gene has already been designated as *MTS2* (for multiple tumour suppressor 2), we favour the name *INK4B*, because it denotes a demonstrated property of the gene product.

p15^{INK4B} was first noted in immunoprecipitates from HaCaT cells which had been arrested in G1 by serum starvation and subsequent restimulation in the presence of TGF- β . By comparison, asynchronous, rapidly proliferating HaCaT cells contained considerably lower levels of p15 in CDK4 and CDK6 immunoprecipitates (Fig. 3a). To separate effects of TGF- β from effects of G1 arrest, asynchronous cultures were treated with TGF- β for various periods, after which patterns of CDK4- and CDK6-associated proteins were examined. Although CDK-associated p16 levels were unaffected, TGF- β treatment increased association of p15 with CDK4 and CDK6 (Fig. 3a). Significant increases in p15 levels were observed within 2–4 h after treatment, and peak levels were reached after 6–8 h (Fig. 3a). Notably, TGF- β treatment also resulted in loss of CDK6 kinase activity from HaCaT cell lysates with a time course that paralleled p15 induction (Fig. 3b).

* To whom correspondence should be addressed.

Two other mechanisms for TGF- β -mediated cell cycle arrest have been previously proposed. In mink lung fibroblasts (Mv1Lu cells), TGF- β treatment suppressed CDK4 synthesis¹¹. This was deemed causal as cells could be rendered TGF- β resistant by constitutive overexpression of CDK4. In HaCaT cells, TGF- β treatment had no effect on CDK4 protein (Fig. 3a) or



absence of TGF- β using RNAZOL B according to the manufacturer's instructions. Messenger RNA preparation and cDNA synthesis were exactly as previously described²⁷. Double-stranded cDNA was ligated into λ -ZapII arms according to the manufacturers instructions. Low-stringency hybridization was performed at 50 °C in 500 mM NaPO₄, pH7.0, 1 mM EDTA, 15% formamide, 7% SDS, 0.1% bovine serum albumin (wash: 1 $\times$ SSC, 50 °C). The p15^{INK4B} cDNA was also isolated from a human mammary epithelial cell library provided by M. Stampfer (University of California at Berkeley). Cell lysis, immunoprecipitation and protease mapping were performed exactly as previously described²⁸. For chymotrypsin digestion, either 7 μ g or 1.5 μ g of enzyme was used. Gels for chymotrypsin mapping contained 0.1% SDS. The CDK6 antibody was a gift of M. Meverson (MGH, Boston)⁴.

FIG. 2 Specific inhibition of CDK4 and CDK6 by p15. **a**, Binding of CDK4 and CDK6 by p15. A fusion protein consisting of p15 and glutathione-S-transferase (GST-p15) was purified from bacteria and mixed with equal quantities of *in vitro* translated, 35 S-labelled CDC2, CDK2, CDK4, CDK5 or CDK6 (as indicated). Bound proteins were recovered on glutathione sepharose and then released by boiling in SDS gel sample buffer. Bound proteins were analysed by electrophoresis in a 17.5% PAGE gel. For comparison, a similar experiment using GST-p16 is also shown. **b**, p15 binding inhibits cyclin D-CDK4 kinase. Active cyclin D-CDK4 kinase, present in lysates of baculovirus-infected insect cells, was incubated with increasing quantities of GST-p15^{INK4B} for 30 min at 30 °C. Cyclin D-CDK4 kinase activity was then assayed using bacterially produced GST-Rb and [γ - 32 P]ATP as substrates. For comparison, a parallel experiment using GST-p16^{INK4} is shown. For all experiments shown in this figure, quantities of intact GST-p15 and GST-p16 were similar. **c**, p15 inhibits cyclin D-CDK6 kinase. Cyclin D-CDK6 complexes were produced in baculovirus-infected insect cells and incubated with increasing quantities of GST-p15 or GST-p16 (indicated) as described above. Following incubation, the ability of these complexes to catalyse GST-Rb phosphorylation was determined. For **b** and **c**, GST-Rb was recovered on glutathione sepharose and released by boiling in SDS sample buffer before electrophoresis.

METHODS. Preparation of GST-p16 fusion protein from bacteria has been described previously⁹. GST-p15 was prepared identically. For the binding assays *in vitro* translated CDKs were prepared using the TNT-lysate *in vitro* translation kit (Promega) according to the manufacturer's instructions. For the binding assay, GST-p15 or GST-p16 (250 ng) was incubated with *in vitro* translated CDKs for 30 min at 30 °C in 30 μ l containing 20 mM Tris, pH 8.0, 10 mM MgCl₂, 1 mM EGTA. Following incubation, mixtures were diluted to 250 μ l in IP buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 0.5% NP-40) and incubated for 1 h with 12.5 μ l

of glutathione sepharose. Beads were washed four times with 1 ml of IP buffer before release of bound proteins. For kinase assays baculovirus lysates containing active cyclin D-CDK4 were prepared as described previously^{9,28}. Lysates containing active cyclin D-CDK6 were prepared identically (the CDK6 baculovirus was a gift of M. Meyerson, MGH). For kinase assays 10 μ l of baculoviral lysates were mixed with approximately 0, 10 ng, 20 ng, 50 ng, or 100 ng GST-p15 or GST-p16 and incubated for 30 min at 30 °C in a total volume of 30 μ l containing 20 mM Tris, pH 8.0, 10 mM MgCl₂, 1 mM EGTA. Following incubation, 0.5 μ g GST-Rb and 0.5 μ l [γ - 32 P]ATP (5 μ Ci, 3000 Ci mmol⁻¹) were added for 10 min at 30 °C. Reactions were stopped by the addition of 250 μ l IP buffer and 15 μ l glutathione sepharose. After a further hour of incubation at 4 °C, beads were washed four times with IP buffer before release of bound proteins and subsequent electrophoresis.

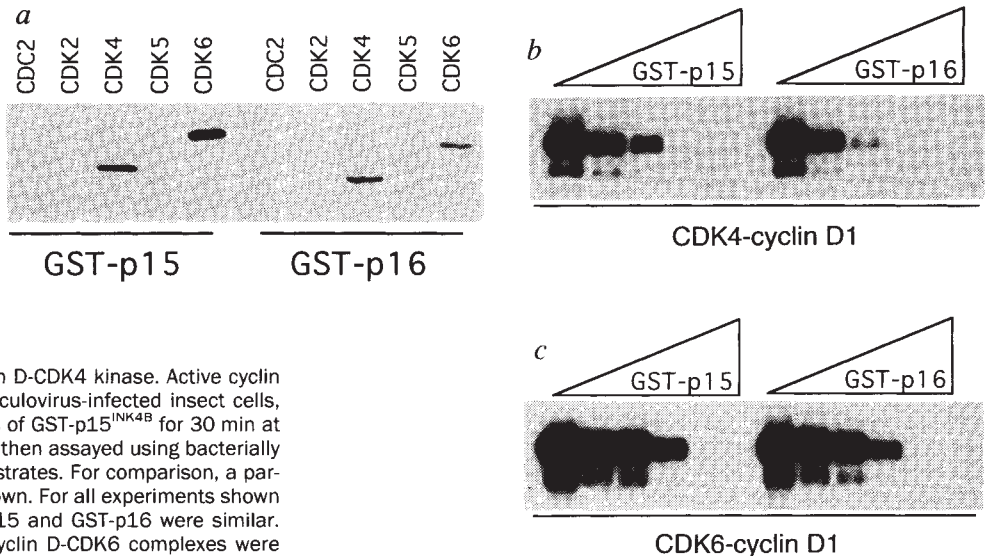
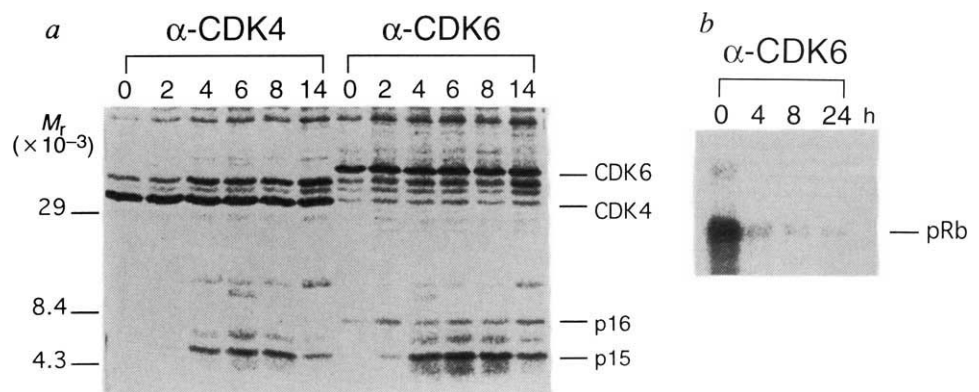


FIG. 3 **a**, Increased association of p15^{INK4B} with CDK4 and CDK6 after TGF- β treatment. Asynchronous HaCaT cells at ~50% confluence were treated with 2 ng ml⁻¹ TGF- β for the times indicated. For the last 2 h of TGF- β treatment, cells were labelled with [35 S]methionine. Cells were lysed and proteins were immunoprecipitated using either α -CDK4 or α -CDK6 antibodies (as indicated). The positions of p15, p16, CDK4 and CDK6 are marked. The positions of labelled protein molecular weight markers electrophoresed in parallel are also shown. **b**, Inhibition of CDK6 kinase after TGF- β treatment. Asynchronous cells were treated for the indicated times with TGF- β (2 ng ml⁻¹).

Cells were lysed in IP buffer (see above) and proteins were recovered by immunoprecipitation with a CDK6 antibody (gift of M. Meyerson). The kinase activity of CDK6 complexes present in the immunoprecipitates was tested using GST-Rb as a substrate⁴. Parallel immunoprecipitates, analysed for protein composition of CDK6 complexes, showed significant accumulation of p15 starting at the 4-h time point (data not shown). The position of the phosphorylated pRb is indicated.

METHODS. HaCaT cells were cultured as described above. TGF- β treatment was initiated by adding TGF- β to existing media to 2 ng ml⁻¹. For cell labelling, medium was changed to DMEM minus methionine (Gibco-BRL) containing 0.2 ng ml⁻¹ TGF- β , 10% dialysed FBS and 0.5 mCi ml⁻¹ [35 S]-methionine/cysteine (trans-label, NEN). Cell lysis and immunoprecipitation were as described above. HaCaT cultures treated with TGF- β in parallel were stained with DAPI and analysed by FACS. For 8 h following TGF- β addition, there was no appreciable change in the percentage of G1 cells. After 14 h, the G1 population began to increase. Arrest of HaCaT cells in G1 (~85% G1 cells) required at least 24 h after treatment



of an asynchronous culture. For kinase assays, 35 S-labelled, TGF- β -treated cells were lysed in IP buffer (see above). Proteins were recovered using a CDK6-specific antiserum and protein A-sepharose (Pierce). Immunoprecipitates were split for analysis of complex components (not shown) and kinase assays. The beads reserved for the kinase assay were washed three times in IP buffer then twice in kinase buffer (see above). (20 μ l) Kinase buffer containing 0.5 μ g GST-Rb and 5 μ Ci [γ - 32 P]ATP were then added to the beads, and the reactions were incubated for 10 min at 30 °C. Reactions were stopped by the addition of 300 μ l IP buffer, and protein A-Sepharose was removed by centrifugation. Glutathione sepharose (15 μ l) was added to the supernatant to separate GST-Rb from unreacted [γ - 32 P]ATP. Subsequent processing and analysis were as described in the legend to Fig. 2. To date, using our CDK4 and cyclin D1 antibodies, we have been unable to assay CDK4 kinase from a variety of cell types and thus were not able to examine the effects of TGF- β on the activity of these complexes.

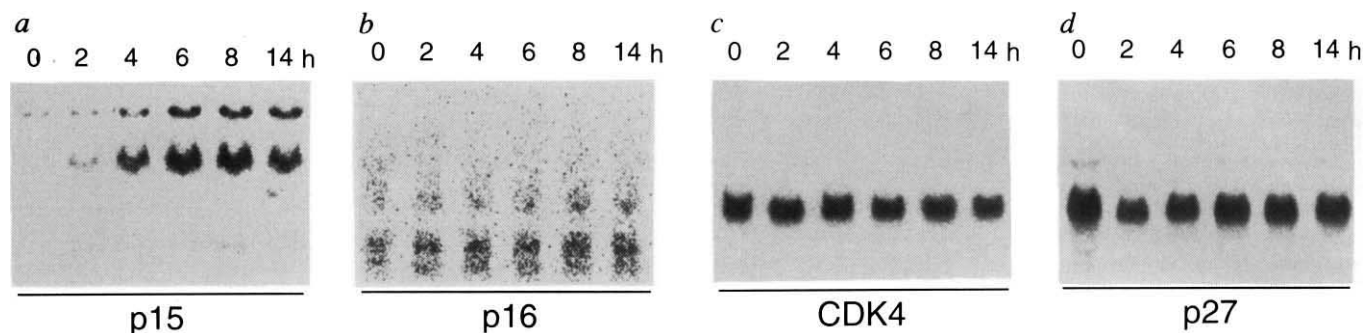
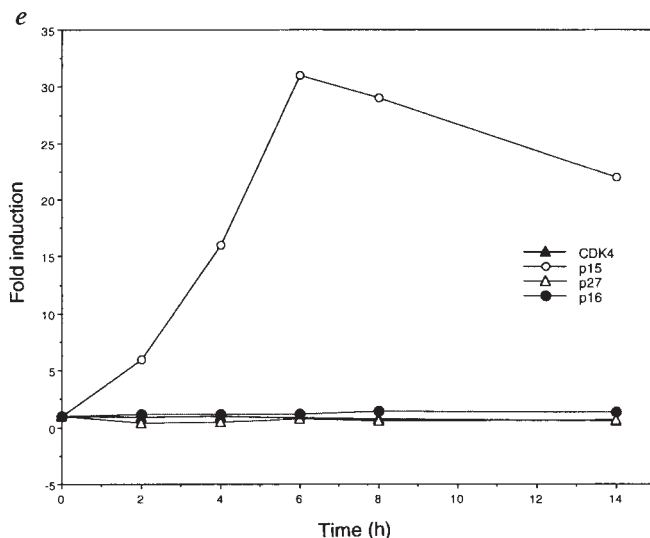


FIG. 4 Induction of p15 mRNA by TGF- β treatment *a*, Asynchronous HaCaT cultures were treated with TGF- β for the indicated times. Total RNA was prepared from treated cells and used for northern blotting with a probe specific for p15^{INK4B}, which recognizes three p15 mRNAs of around 0.8, 2.2 and 3.2 kb. *b*, RNAs identical to those used in *a* were probed with a fragment specific for p16^{INK4}. *c*, RNA from TGF- β -treated HaCaT cells (see *a*) was probed with a fragment encoding human CDK4. *d*, RNA from TGF- β -treated HaCaT cells was probed with a fragment of p27^{Kip1} cDNA (a gift of H. Zhang, Cold Spring Harbor Laboratory). *e*, Northern signals from panels *a*–*d* were quantitated on a Fuji BAS2000 phosphorimager and plotted to give a graphical representation of the results. In each case, the amount of transcript detected in the absence of TGF- β is defined as 1.0. Relative increases or decreases in mRNA abundance are plotted with respect to the zero time point.

METHODS. HaCaT cells were treated with TGF- β as described in the legend to Fig. 3. RNA was prepared from treated cultures using RNeasy Lysate B according to the manufacturer's instructions. p15 and p16-specific probes consisted of the first coding exons of these genes and were prepared by PCR. No cross hybridization of probes was evident under the conditions used for these northern blots (hybridization: 200 mM NaPO₄, pH 7.0, 1 mM EDTA, 15% formamide, 7% SDS, 0.1% bovine serum albumin at 65 °C; wash: 0.2 × SSC, 65 °C). The CDK4 and p27^{Kip1} probes consisted of the coding sequences of the respective cDNAs. RNA amounts were normalized by mass, and over-probing of the blot



shown in *c* with a human actin probe suggested no more than 10% variance in RNA amounts between lanes.

mRNA (Fig. 4c, e) levels. On the basis of the properties of p15, we would predict that CDK4 overexpression could also render HaCaT cells TGF- β resistant by titrating the p15 CDK4/CDK6 inhibitor. p27^{Kip1}, a CDK inhibitor that was purified from TGF- β -arrested cells, has also been proposed as a link between TGF- β and cell cycle control^{12–14}. However, in HaCaT cells, TGF- β treatment had no effect on p27 mRNA levels (Fig. 4d, e). p27 mRNA was also unaffected by similar treatment of Mv1Lu cells¹². Thus in these cell lines, any contribution that p27 might make to TGF- β -mediated cell cycle arrest would have to occur by regulation at the post-translational level.

Considered together, our results suggest that p15^{INK4B} may function as an effector of TGF- β -mediated cell cycle arrest through inhibition of CDK4 and CDK6 kinases. At present, we cannot distinguish whether p15 acts as the sole mediator of TGF- β -induced arrest in some cells or whether p15 must cooperate with other TGF- β -responsive pathways^{1,11–14}. In specific cell types, TGF- β can also induce differentiation¹. Withdrawal from the cell cycle has been proposed as a component of differentiation, and induction of p15 by TGF- β could contribute to this process.

Cytogenetic abnormalities at 9p21 are common in many types of human tumours, suggesting the presence of a tumour suppressor gene at this locus^{15–17}. An inherited cancer syndrome which causes predisposition to melanoma also maps to 9p21 (refs 18–20). p16^{INK4} was initially proposed as a candidate for both of these activities on the basis of analysis of p16 deletions and point mutations in cell lines^{10,21}. However, the presence of a second functional member of the p16 family at 9p21 raises the possibility that loss of tumour suppression may involve inactivation of either or both genes. The response of p16 to viral oncoproteins indicates that it may function in intracellular growth

regulatory pathways^{9,22}, and results presented here suggest that p15 may act as an effector of extracellular growth inhibitory signals. Thus, deletions of 9p21 which remove both genes (or other mutations that might inactivate both) could simultaneously negate two major proliferation control pathways. In this regard, the ability of TGF- β to induce growth arrest is reduced or lost in many neoplastically transformed cell lines^{1,23–25}. In particular, melanocytes are sensitive to growth inhibition by TGF- β , but many metastatic melanoma cells are TGF- β resistant²⁶. □

Received 7 July; accepted 29 July 1994.

1. Massague, J. A. *Rev. Cell Biol.* **6**, 597–641 (1990).
2. Sherr, C. J. *Cell* **73**, 1059–1065 (1993).
3. Bates, S. et al. *Oncogene* **9**, 71–79 (1994).
4. Meyerson, M. & Harlow, E. *Molec. cell. Biol.* **14**, 2077–2086 (1994).
5. Laiho, M., DeCaprio, J. A., Ludlow, J. W., Livingston, D. M. & Massague, J. *Cell* **62**, 175–185 (1990).
6. Pietenpol, J. et al. *Cell* **61**, 777–785 (1990).
7. Boukamp, P. et al. *J. Cell Biol.* **106**, 761–771 (1988).
8. Geng, Y. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10315–10319 (1993).
9. Serrano, M., Hannon, G. J. & Beach, D. *Nature* **366**, 704–707 (1993).
10. Kamb, A. et al. *Science* **264**, 436–440 (1994).
11. Ewen, M. E., Sluss, H. K., Whitehouse, L. L. & Livingston, D. M. *Cell* **74**, 1009–1020 (1993).
12. Polyak, K. et al. *Cell* **78**, 59–66 (1994).
13. Toyoshima, H. & Hunter, T. *Cell* **78**, 67–74 (1994).
14. Polyak, K. et al. *Genes Dev.* **8**, 9–22 (1994).
15. Bello, M. J. et al. *Genes Chromosomes Cancer* **9**, 33–41 (1994).
16. Merlo, A., Gabrielson, E., Askin, F. & Sidransky, D. *Cancer Res.* **54**, 640–642 (1994).
17. Cheng, J. Q., Jhanwar, S. C., Lu, Y. Y. & Testa, J. R. *Cancer Res.* **53**, 4761–4763 (1993).
18. Petty, E. M. et al. *Am. J. hum. Genet.* **53**, 96–104 (1993).
19. Fountain, J. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10557–10561 (1992).
20. Cannon-Albright, L. A. et al. *Science* **258**, 1148–1152 (1992).
21. Nobori, T. et al. *Nature* **368**, 753–756 (1994).
22. Xiong, Y., Zhang, H. & Beach, D. *Genes Dev.* **7**, 1572–1583 (1993).
23. Moses, H. L. et al. in *Cancer Cells* (eds Feramisco, J., Ozanne, B. & Stiles, C.) **3**, 65–71 (1985).
24. Masui, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2438–2442 (1986).

25. Manning, A. M., Williams, A. C., Gamie, S. M. & Paraskeva, C. *Oncogene* **6**, 1471–1476 (1991).
 26. Rodeck, U. et al. *Cancer Res.* **54**, 575–581 (1994).
 27. Hannon, G. J., Demetrick, D. & Beach, D. & Beach, D. *Genes Dev.* **7**, 2378–2391 (1993).
 28. Xiong, Y. et al. *Nature* **366**, 701–704 (1993).

ACKNOWLEDGEMENTS. We thank S. Matsumoto and C. Gawel for technical assistance and P. Renna, M. Ockler and J. Duffy for photography and graphics. M. Stampfer (Berkeley), M. Meyerson (MGH), P. Boukamp and N. Fusenig (German Cancer Research Center) provided critical reagents. We are grateful to Hui Zhang for providing baculoviral lysates and for helpful discussions. G.J.H. is a fellow of the Damon Runyon–Walter Winchell Cancer Research Fund. D.B. is an Investigator of the Howard Hughes Medical Institute. This work was supported in part by the NIH.

Location of a folding protein and shape changes in GroEL–GroES complexes imaged by cryo-electron microscopy

Shaoxia Chen*, Alan M. Roseman*, Allison S. Hunter*, Stephen P. Wood*, Steven G. Burston†, Neil A. Ranson†, Anthony R. Clarke† & Helen R. Saibil*

* Department of Crystallography, Birkbeck College London, London WC1E 7HX, UK

† Department of Biochemistry and Centre for Molecular Recognition, University of Bristol, Bristol BS8 1TD, UK

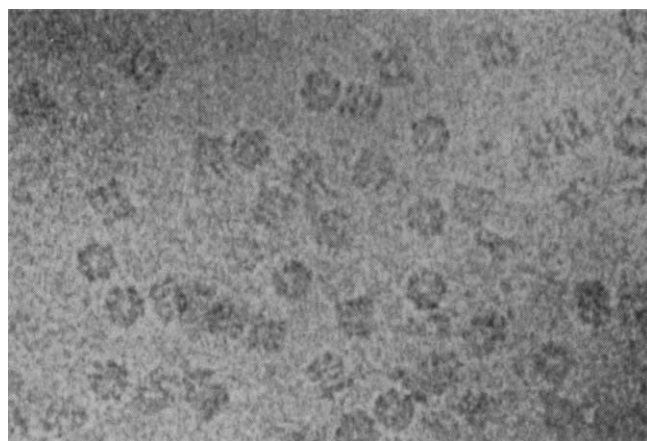
PROTEIN folding mediated by the molecular chaperone GroEL occurs by its binding to non-native polypeptide substrates and is driven by ATP hydrolysis¹. Both of these processes are influenced by the reversible association of the co-protein, GroES (refs 2–4). GroEL and other chaperonin 60 molecules⁵ are large, cylindrical oligomers consisting of two stacked heptameric rings of subunits^{6,7}; each ring forms a cage-like structure⁸ thought to bind polypeptides in a central cavity^{8–10}. Chaperonins play a passive role in folding by binding or sequestering folding proteins to prevent their aggregation^{11–13}, but they may also actively unfold substrate proteins trapped in misfolded forms, enabling them to assume productive folding conformations^{14–16}. Biochemical studies show that GroES improves the efficiency of GroEL function^{2,3,17}, but the structural basis for this is unknown. Here we report the first direct visualization, by cryo-electron microscopy, of a non-native protein substrate (malate dehydrogenase) bound to the mobile, outer domains at one end of GroEL. Addition of GroES to GroEL in the presence of ATP causes a dramatic hinge opening of about 60°. GroES binds to the equivalent surface of the GroEL outer domains, but on the opposite end of the GroEL oligomer to the protein substrate.

Size-exclusion chromatography of purified GroEL in the presence of ATP releases a significant amount of residually bound protein and peptide to generate an 'emptied' state. Cryo-electron microscope (cryo-EM) side- and end-view projections (Figs 1 and 2a, b) of this emptied GroEL were aligned and averaged, and a three-dimensional reconstruction was calculated from the averages by backprojecting the seven symmetry-related side views and the end view. Figure 1c shows a central section (21 Å thick) of this map. The side projection shows that the overall dimensions of the 14-mer are 140 Å high by 130 Å wide. The end projection is not 7-fold averaged, and demonstrates the strong 7-fold symmetry of the structure. The side-view section (Fig. 2c) shows the two major domains of each subunit, with the inner domains forming the interface between top and bottom heptameric rings, and outer domains surrounding a 35–40-Å diameter hole. A three-dimensional surface-contoured view of the structure is shown in Fig. 4b.

The stable complex between GroEL and malate dehydrogenase (MDH) was formed by mixing GroEL with a 2.4-fold

molar excess of chemically denatured MDH (GroEL 14-mer : MDH monomer) to saturate the available binding sites (see Fig. 2 legend). Figure 2d–f shows the cryo-EM images of this binary complex. Comparison with the 'empty' structure reveals a distinct region of additional density at only one end of the oligomer in the side views, and extra density fills the hole seen in the end projection. Difference maps of the end projection and the section are shown in Fig. 2g and h with the empty GroEL contours superimposed. These maps locate the MDH density in one end cavity of GroEL, between the outer domains and protruding slightly. The visible ligand density (which has been rotationally averaged about the vertical axis in the section view; Fig. 2h) occupies a volume roughly similar to that of the 34,000-*M_r* native subunit (approximately 50 × 30 × 35 Å), but diffuse density may extend into a larger volume. The asymmetric complex is consistent with a spectroscopically determined binding stoichiometry of 1:1 (refs 18, 19, and S.G.B. and A.R.C., unpublished observations), and implies negative cooperativity of ligand binding.

Binding of Mg-ATP to GroEL is sufficient to release some substrate proteins, but others are not released into productive folding forms by ATP unless GroES is also present. To investigate the structural influence of Mg-ATP alone, a set of cryo-EM images of empty GroEL incubated in Mg-ATP were collected and are shown in Fig. 3a–c. Compared with GroEL (Fig. 2c), GroEL-ATP shows a slight vertical expansion of the oligomer (from 140 to 150 Å) owing to an opening of the outer domains (Fig. 3a, c). This outer domain displacement differs from that deduced from previous negative stain electron microscopic work⁸. An important factor seems to be the presence of tightly bound substrates in the earlier study, as cryo-EM images of the GroEL-ATP complex made from non-emptied GroEL more closely resemble the previous negative stain findings (not shown). On the basis of the present cryo-EM data, there may be a slight asymmetry in the GroEL-ATP complex. The three-dimensional view of GroEL-ATP in Fig. 4c shows the petal-like opening of the outer domains.



200 Å

FIG. 1 Unstained, frozen-hydrated GroEL oligomers were imaged in vitreous ice over holes in the carbon support film. The dark regions represent protein density, contrasted against the background of vitreous ice. Two orientations of the oligomers are present: end views, which show the 7-fold symmetry, and side views with four layers of density. METHODS. Solutions containing 1–2 mg ml⁻¹ GroEL in 10 mM ammonium acetate and 10 mM Tris buffer pH 7.6 were prepared at room temperature, vitrified by plunging into liquid ethane and imaged at -170 °C. Cryo-EM was done on a JEOL 1200 EX microscope with an Oxford Instruments cryotransfer stage. All images used in the following analysis were recorded with an electron dose ≈ 10 e⁻ Å⁻² and with a defocus in the range 500–850 nm.

The vertical expansion of GroEL induced by nucleotide binding is greatly enhanced in the presence of GroES. In the conditions used here, the great majority of complexes formed between GroEL and GroES in the presence of ATP are asymmetrical, with GroES bound to one end of GroEL (Fig. 3*d-f*). The interaction with GroES traps the adjacent GroEL outer domains in a wide-open conformation (Fig. 3*f*). Estimating orientations by taking lines along the direction of elongation of the density, we find that the angle between inner and outer domains increases by as much as 60°. A smaller domain movement brought about by a 5–10° rotation of the hinge occurs on binding ATP alone (Fig. 3*c*), and also in the ring not bound to GroES in the complex (Fig. 3*f*, bottom). An enclosed, dome-shaped volume of maximum height 65 Å and maximum width 80 Å is formed by GroES binding and the hinge opening. A surface-contoured view of the three-dimensional map of the complex is shown in Fig. 4*d*.

When the stable GroEL–MDH complex is combined with GroES and ATP, MDH folding is initiated (N. Dunster and A.R.C., unpublished results). To determine the relationship

between GroES and substrate-binding sites, we trapped this folding complex by vitrification 15 s after adding GroES and ATP (Fig. 3*g-i*). The MDH density occurs at the opposite end of GroEL to GroES, between the free outer domains, which are slightly farther apart, as in the ATP form. A similar ternary complex is seen with ADP as the nucleotide (not shown). Note that after 15 s approximately half the initially bound ATP would have been hydrolysed^{14,20}.

Figure 4*a* is a diagram representing the subunit outlines of the chaperonin complexes. The ends of the GroEL outer domains provide the binding sites for the protein substrate, and these surfaces are pushed apart on binding ATP as a consequence of the rotation about the inter-domain hinge. When ATP is present, GroES binds to the equivalent surface on the widely opened outer domains, creating the enclosed cavity in the GroEL–GroES complex. In the ternary complex, substrate is not seen in the GroES-enclosed cavity, but is found in the binding cavity on the opposite ring.

There is also evidence for substrate binding to the side of GroEL opposite GroES by antibody labelling²¹, but a cross-

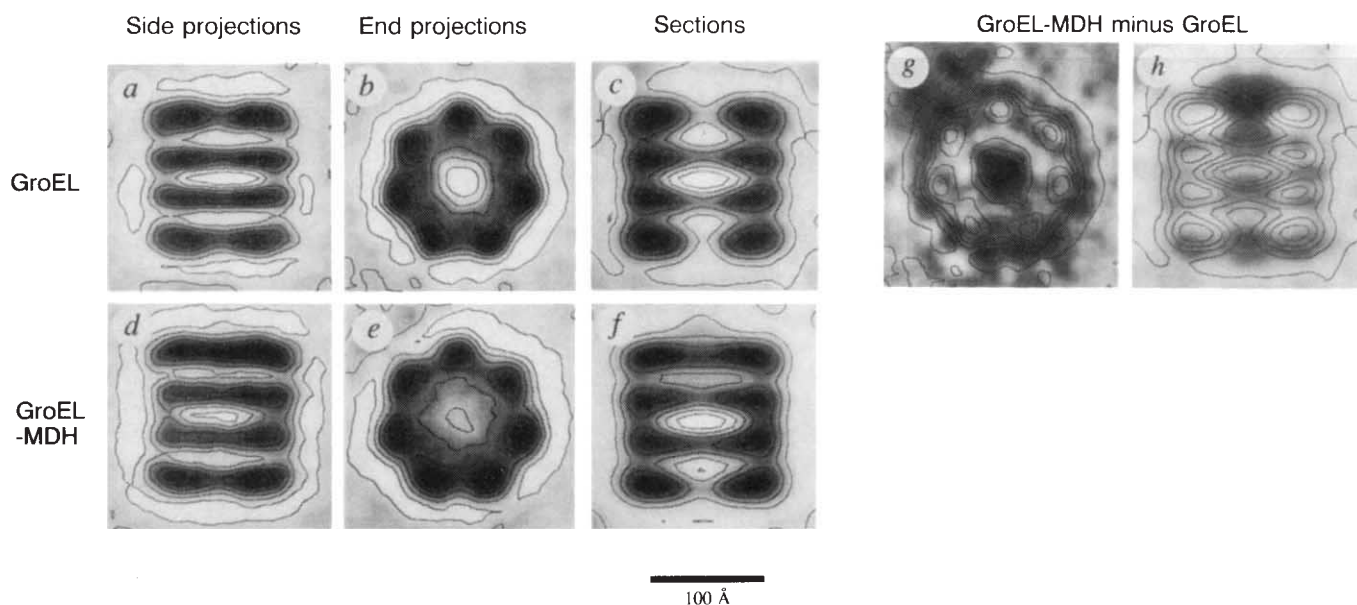


FIG. 2 *a*, An averaged side-view projection of 418 individual images aligned by correlation methods, showing the typical four-layered density. *b*, End-view projection, average of 490 molecules, showing the 7-fold symmetry. *c*, A 21-Å-thick section through the 3-D map generated from *a* and *b*, showing four subunits in vertical section. The connecting bridges between the domains and also the subunit contacts between rings are near the outer surface of the cylinder. *d-f*, The corresponding views of the GroEL–MDH complex. The side projection (*d*, average of 368 views) and section *f* show additional density at one end of the oligomer, and the end projection (*e*, average of 329 views) shows the extra density in the central hole. *g, h*, Difference maps of GroEL–MDH minus GroEL, showing the distribution of MDH density in the complex. The differences are shown as grey-scale density, with the contours of empty GroEL superimposed for reference. The darkest regions show the location of MDH density in the binary complex. *g*, End-view difference map, showing the MDH density in the central hole in projection. No symmetry averaging was applied to this view. *h*, Section difference map, showing the location of rotationally averaged MDH between the outer domains of GroEL and protruding slightly from the GroEL surface. The vertical position of MDH is not affected by the symmetry averaging.

METHODS. *Escherichia coli* GroEL and GroES were prepared as described²⁷. To discharge residual bound polypeptide, GroEL was applied to a Sepharose 4-B chromatography column (100 cm × 2.6 cm) in 2 mM ATP in 20 mM MgCl₂, 50 mM KCl, 50 mM triethanolamine-

HCl, pH 7.5. The GroEL–MDH complex was prepared by rapidly injecting 200 µl of 1.2 µM GroEL (14-mer) onto a 2.8-µl drop of 200 µM porcine mitochondrial MDH (subunits) in 3 M guanidinium chloride. Circular dichroism spectra showed complete loss of MDH secondary structure at 2 M guanidinium chloride. Saturation of MDH binding at 1:1 stoichiometry was confirmed by two methods, that is, suppression of recovery of enzyme activity and titration of binding of pyrene-labelled, unfolded MDH to GroEL. Micrographs were recorded at ×30,000 on Agfa EM film and digitized with a CCD camera at a final resolution of 5.6 Å per pixel. Particle alignment was done by iterative cross-correlation, using procedures written for SPIDER and SEMPER software systems. Averages of manually aligned views for each sample provided initial templates for cross-correlation. To assess the resolution, the raw images forming each average were split to form two half-averages, and phase residuals were determined by comparing the two halves, showing that the information was significant to ~25 Å. Three-dimensional reconstructions calculated by back-projection from symmetry-related, averaged views had sufficient angular resolution (indicating preferential orientation about the 7-fold axis parallel to the air–water interface) to align them with the corresponding end projections. Inclusion of the end views in the reconstructions improved the angular resolution but had little effect on the appearance of the 21-Å-thick section views. Surface-contoured representations of the 3-D maps are shown in Fig. 4*b–d*.

linking study implies substrate binding adjacent to GroES²². The 'same-side' complex may be present transiently, as GroES is bound and released during cycles of ATP binding and hydrolysis², and GroES can bind to both ends simultaneously^{20,23-25}. However, the double-sided complex is not symmetrical, as the second GroES binds more weakly and the nucleotide state may differ at the two ends²⁰. Given that the substrate can remain bound to the ATP ternary complex for

≥ 15 s and if the GroES-GroEL-GroES complex occurs as an intermediate, there remains the possibility of a transient state in which GroES encapsulates the substrate. However, we have no evidence for occupation of a GroES-capped cavity, and, furthermore, our results cast doubt on the recent proposal that substrate proteins bind to the exterior of the GroEL cylinder²⁶.

This work gives the first direct image of a chaperone-bound, folding protein and allows visualization of dramatic structural

FIG. 3 Cryo-electron microscopy of GroEL-ATP (a-c), GroEL-GroES-ATP complexes (d-f) and GroEL-GroES-ATP-MDH complexes (g-i). a, Side projection average of 186 views of GroEL-ATP. The overall size of the oligomer is 150×130 Å. b, End projection average of 496 views. c, Section of the 3-D map generated from a and b, as in Fig. 1, showing the enlarged gap between the outer domains. d, Side projection average of 245 views, showing GroES binding to one end of GroEL, which is in a more open configuration. e, End projection average of 252 views, showing the 7-fold symmetry. f, Section of 3-D map generated from d and e, showing the large hinge opening between GroEL subunit domains in the ring binding GroES, and the internal cavity formed by GroES capping the opened-up structure. g, Side projection average of 328 views of the GroEL-MDH-GroES-ATP ternary complex, showing additional density in the ring opposite GroES. h, Average of 243 end views of the ternary complex. i, Section of the 3-D map generated from g and h, showing MDH density in the lower ring. METHODS. The GroEL-ATP and GroEL-GroES-ATP complex were formed by mixing $1.2 \mu\text{M}$ GroEL (14-mer) with 2 mM ATP in the standard buffer, with or without $3.6 \mu\text{M}$ GroES (7-mer). The ternary complex was formed by adding GroES and ATP to the GroEL-MDH complex formed as described in Fig. 2 legend. The GroEL-GroES complexes were frozen 15–30 s after mixing. GroEL-ATP samples were frozen after 10–20 min of incubation in Mg-ATP, and the ternary complex samples were frozen 15 s after adding GroES and ATP. Addition of 40 mM guanidinium chloride alone had no effect on the appearance of the GroEL-GroES-ATP complex, and very similar ternary complexes were obtained with acid- and urea-denatured MDH (not shown).

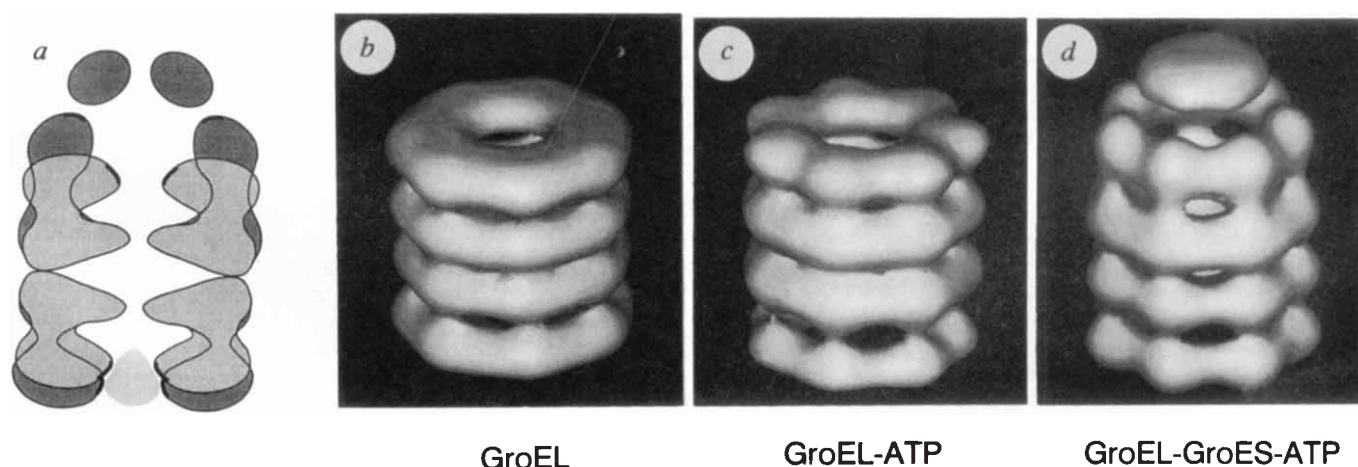
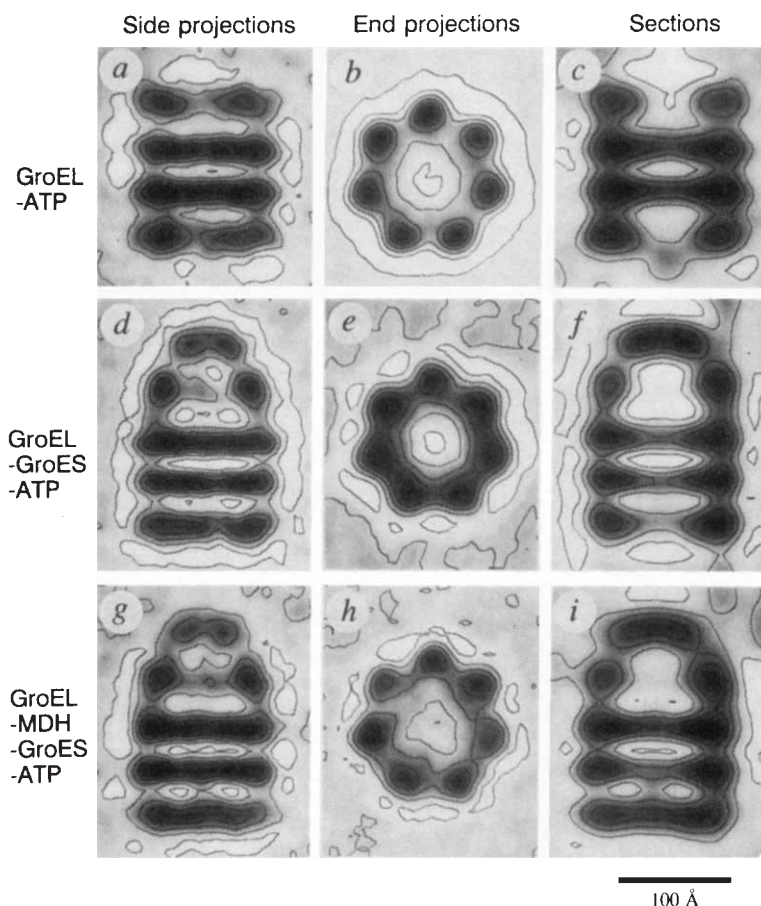


FIG. 4 a, Diagram derived from the outer contours of empty GroEL and GroEL-GroES sections, showing the dramatic reorientation of the outer domains of GroEL in contact with GroES. The ligand position is shown as a shaded area to illustrate the point that the end of the outer domain, which binds substrate, rotates out to interact with GroES. The interaction site is indicated by a darker line. b, The external shape of GroEL seen in a surface-rendered view of the 3-D reconstruction as a solid,

reflective object slightly tipped towards the viewer. c, 3-D reconstruction of GroEL-ATP, showing the opening of the outer domains widening the substrate-binding cavity. d, The same representation of the GroEL-GroES-ATP complex, showing the pronounced vertical expansion of GroEL, the location of GroES as a disk above GroEL, and the enlargement of the cavity. 3-D surface rendering was done on the University of London Computer Centre Convex using the program AVS.

changes associated with ATP and GroES binding. GroEL provides binding surfaces for substrate proteins on a ring of highly mobile domains. □

Received 20 May; accepted 24 August 1994.

- Hendrick, J. P. & Hartl, F. U. *A. Rev. Biochem.* **62**, 349–384 (1993).
- Martin, J., Mayhew, M., Langer, T. & Hartl, F. U. *Nature* **366**, 228–233 (1993).
- Fisher, M. J. *biol. Chem.* **269**, 13629–13636 (1994).
- Schmidt, M., Buchner, J., Todd, M. J., Lorimer, G. H. & Viitanen, P. V. *J. biol. Chem.* **267**, 10304–10311 (1994).
- Kubota, H., Hynes, G., Carne, A., Ashworth, A. & Willison, K. *Curr. Biol.* **4**, 89–99 (1994).
- Hendrix, R. W. *J. molec. Biol.* **129**, 375–392 (1979).
- Hutchinson, E. G., Tichelaar, W., Hofhaus, G., Weiss, H. & Leonard, K. *EMBO J.* **8**, 1485–1490 (1989).
- Saibil, H. R. *et al. Curr. Biol.* **3**, 265–273 (1993).
- Langer, T., Pfeifer, G., Martin, J., Baumeister, W. & Hartl, F. U. *EMBO J.* **11**, 4757–4765 (1992).
- Braig, K., Simon, M., Furuya, F., Hainfeld, J. F. & Horwich, A. L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3978–3982 (1993).
- Ellis, R. J. & Hemmingsen, S. M. *Trends biochem. Sci.* **14**, 339–342 (1989).
- Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. *Nature* **342**, 884–889 (1989).
- Nilsson, B. & Anderson, S. A. *Rev. Microbiol.* **45**, 607–635 (1991).
- Jackson, G. S. *et al. Biochemistry* **32**, 2554–2563 (1993).
- Zahn, R., Spitzfaden, C., Ottiger, M., Wüthrich, K. & Plückthun, A. *Nature* **368**, 261–265 (1994).
- Peralta, D., Hartman, D. J., Hoogenraad, N. J. & Høj, P. B. *FEBS Lett.* **339**, 40–45 (1994).
- Martin, J. *et al. Nature* **352**, 36–42 (1991).
- Mendoza, J. A., Lorimer, G. H. & Horowitz, P. M. *J. biol. Chem.* **266**, 16973–16976 (1991).
- Badcoe, I. G. *et al. Biochemistry* **30**, 9195–9200 (1991).
- Todd, M. J., Viitanen, P. V. & Lorimer, G. H. *Science* **265**, 659–666 (1994).
- Ishii, N., Taguchi, H., Sasabe, H. & Yoshida, M. *J. molec. Biol.* **236**, 691–696 (1994).
- Bochkareva, E. S. & Girshovich, A. S. *J. biol. Chem.* **267**, 25672–25675 (1992).
- Harris, J. R., Plückthun, A. & Zahn, R. *J. struct. Biol.* (in the press).
- Llorca, O., Marco, S., Carrascosa, J. L. & Valpuesta, J. M. *FEBS Lett.* **345**, 181–186 (1994).
- Schmidt, R. *et al. Science* **265**, 656–659 (1994).
- Azem, A., Kessel, M. & Goloubinoff, P. *Science* **265**, 653–656 (1994).
- Stanforth, R. A. *et al. FEBS Lett.* **344**, 129–135 (1994).

ACKNOWLEDGEMENTS. We thank J. Munn for EM support, R. Westlake and J. Bouquiere for computing support, R. J. Ellis and G. Laxer for discussion, the Birkbeck Photo Unit for photographic work and the Wellcome Trust and AFRC for project grants. A.R.C. is a Lister Institute Fellow.

Context is a major determinant of β -sheet propensity

Daniel L. Minor Jr* & Peter S. Kim†

Departments of * Chemistry and † Biology, Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

RESIDUES in β -sheets occur in two distinct tertiary contexts: central strands, bordered on both sides by other β -strands, and edge strands, bordered on only a single side by another β -strand¹. The $\Delta\Delta G$ values for β -sheet formation measured at an edge β -strand of the IgG-binding domain of protein G (GB1) are quite different from those obtained previously^{2,3} at a central position in the same protein. In particular, there is no correlation at the edge position with statistically determined β -sheet-forming preferences⁴. The differences between β -sheet propensities measured at central and edge β -strands, $\Delta\Delta\Delta G$ values, correlate with the values of water/octanol transfer free energies⁵ and side-chain non-polar surface area for the amino acids⁶. These results strongly suggest that, unlike α -helix formation, β -sheet formation is determined in large part by tertiary context, even at solvent-accessible sites, and not by intrinsic secondary structure preferences.

The 20 naturally occurring amino acids were substituted at a solvent-exposed edge β -strand position, residue 44, by site-directed mutagenesis in a host molecule in which local inter-

TABLE 1 Parameters of β -sheet formation

Amino acid	$\Delta\Delta G$ (kcal mol ⁻¹)	T_m (°C)	$K_a/K_a^{\text{AAA-44Thr}}$	$\Delta\Delta\Delta G$ (kcal mol ⁻¹)
Thr	0.83	60.2	1.0	-0.27
Ser	0.63	59.4	3.1	-0.07
Glu	0.31	57.0	1.5	0.30
Val	0.17	56.1	1.5	-0.65
Phe	0.16	56.1	1.1	-0.70
Tyr	0.11	55.9	1.1	-0.75
Cys	0.08	55.1	0.7	-0.44
Gln	0.04	54.7	1.9	-0.19
Ile	0.02	54.8	1.6	-0.98
Ala	0	54.7	2.9	0
His	-0.01	54.6	1.8	0.01
Met	-0.02	54.2	1.8	-0.74
Asp	-0.10	53.7	1.2	0.84
Trp	-0.17	53.3	3.0	-0.77
Asn	-0.24	52.6	1.0	-0.16
Leu	-0.24	52.8	1.0	-0.75
Lys	-0.40	51.4	0.9	-0.67
Arg	-0.43	51.2	1.0	-0.88
Gly	-0.85	47.6	1.4	0.35
Pro	< -4	< 0	0	—

Listed are $\Delta\Delta G$ values for β -sheet formation at an edge position relative to alanine, thermal melting temperatures (T_m) for the AAA-44Xaa proteins, relative binding constants (K_a) to human Fc for the AAA-44Xaa proteins, and $\Delta\Delta\Delta G$ values for β -sheet formation, comparing $\Delta\Delta G_{\text{edge}} - \Delta\Delta G_{\text{centre}}$ values (see also ref. 2). The K_a for wild-type GB1 is $1.4 \times 10^8 \text{ M}^{-1}$ (ref. 30). The $K_a/K_a^{\text{AAA-44Thr}}$ value for wild-type GB1 is 7.1. ΔG values for unfolding at 321 K were calculated as described previously² using data obtained from CD thermal unfolding measurements and the Gibbs-Helmholtz equation. A positive $\Delta\Delta G$ value indicates an increase in stability relative to alanine. Estimated errors in determination of T_m and ΔG are $\pm 0.5^\circ \text{C}$ and $\pm 0.06 \text{ kcal mol}^{-1}$ respectively. Fc binding was measured as described previously² by adding variable amounts of competitor protein to a fixed quantity of protein G-alkaline phosphatase at 5°C in 96-well plates. The ratio of affinity constants for the mutants was normalized to AAA-44Thr. GB1-Thr1 (see Fig. 1 legend) was included as an internal standard on each plate.

actions to the guest site had been minimized by replacing the nearest neighbours with alanine (see Fig. 1 legend). This edge β -strand is bordered on one side by another β -strand and on the other side by solvent (Fig. 1a). The stability of each protein (denoted AAA-44Xaa) was measured by thermal unfolding as monitored by circular dichroism (CD) at 218 nm (Fig. 1b). $\Delta\Delta G$ values for β -sheet formation, referenced to alanine, were obtained by assuming that changes in global stability result entirely from changes in the ability of the residue at the guest site to adopt a β -sheet conformation. In support of this assumption, molecules representative of the entire $\Delta\Delta G$ range were found to have $\Delta H_{\text{van't Hoff}}/\Delta H_{\text{cal}}$ ratios near unity, indicating that the two-state nature of the equilibrium unfolding transition observed for wild-type GB1 (ref. 7) remains intact (see Fig. 1 legend). The relative free energy differences for β -sheet formation at the edge position are listed in Table 1.

All the proteins tested, with the exception of unfolded AAA-44Pro, bind Fc with around sevenfold reduced affinity relative to wild-type GB1 (Table 1). There is some variation in binding constants but this is not correlated with protein stability. As residues 42–46 have been identified as participants in the Fc-binding interface^{8,9} it seems likely that the observed affinity differences reflect direct effects on binding by substitutions at residue 44.

As a more detailed check on the conformation at the guest site of each molecule, the chemical shifts of the aromatic ring protons of Trp 43 were measured in each of the 20 variants. Trp 43 is expected to be sensitive to changes in structure as it immediately precedes the guest site and is part of the hydrophobic core of the molecule. The chemical shifts of the Trp 43 ring protons are similar in all the variants, with the exception

of unfolded AAA-44Pro, and are significantly different from the chemical shifts for free tryptophan (see Fig. 2 legend). Additionally, unambiguous cross-strand NOEs between the guest strand and its neighbouring strand can be found between upfield shifted $H_{54}^{\gamma 1}$ and $H_{54}^{\gamma 2}$ protons of Val 54 and the H_{43}^{δ} , $H_{43}^{\zeta 2}$, $H_{43}^{\zeta 3}$, $H_{43}^{\epsilon 3}$, $H_{43}^{\eta 2}$, $H_{43}^{\eta \epsilon}$ ring protons of Trp 43 in each folded protein.

Two molecules with substantially different stability—AAA-44Thr and AAA-44Ala—which also have different Fc-binding affinities, were characterized further by NMR. The 1H - ^{15}N correlation spectra of these two molecules are very similar except for resonances from residues in the immediate vicinity of the guest site (Fig. 2a, b). The NMR spectra of both proteins also contain cross-strand nuclear Overhauser effect (NOE) patterns near the guest site that are present in wild-type GB1 (Fig. 2c). Taken together with the Fc-binding and Trp 43 chemical shift results, these data indicate that any structural changes between the variants are small and do not involve significant disruption of the backbone β -sheet structure.

The dramatic effect that context can have on β -sheet propensities is seen in the comparison of the thermodynamic preferences for β -sheet formation measured at the central and edge β -sheet positions (Fig. 3a). Although the overall magnitude of the scale for β -sheet formation at the edge position (~ 2 kcal mol $^{-1}$,

excluding proline) is similar to that measured at the central β -sheet position 2,3 , there is no apparent correlation with the statistically measured β -sheet frequencies of Chou and Fasman 4 (Fig. 3b). In addition, unlike the results obtained at the central position, for which the $\Delta\Delta G$ values are distributed fairly evenly over a wide range, more than half (13/20) of the $\Delta\Delta G$ values measured at the edge position fall within a small range (~ 0.4 kcal mol $^{-1}$).

The context dependence of β -sheet formation suggests that two components contribute to β -sheet propensity: (1) intrinsic ability to form a local extended β -strand structure; and (2) ability to interact with the surrounding tertiary β -sheet structure. The small range of β -sheet propensities measured at the edge position, together with the relative lack of preference for particular side chain rotamers in β -strand residues 10 , suggests that (1) has only a minor role in determining β -sheet propensity. This situation contrasts with α -helix formation where a significantly biased side chain rotamer distribution 10 and a well-distributed range of α -helix propensity values are found $^{11-18}$.

The difference in β -sheet propensity between the edge and centre sites, designated as $\Delta\Delta G = \Delta\Delta G_{edge}^{Xaa} - \Delta\Delta G_{centre}^{Xaa}$, correlates with both the free energy of transfer for the amino acids from octanol to water 5 and with the non-polar accessible

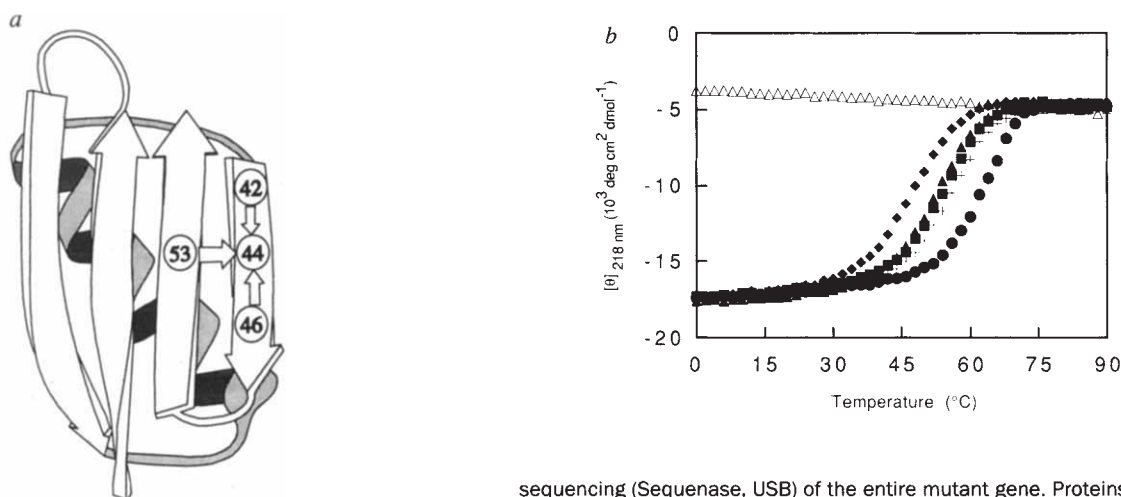
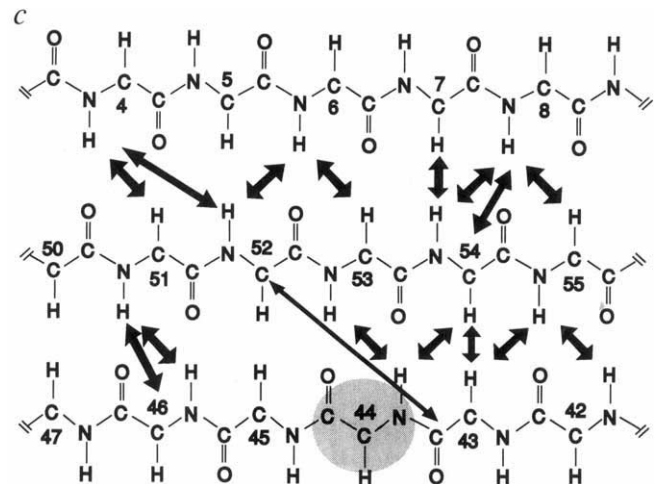


FIG. 1 a, Ribbon drawing 21 of GB1 based on the coordinate set 2GB1 22 . The positions of the guest site and the surrounding residues are indicated. The arrows represent the inter-residue contacts to the guest site made by the surrounding residues. b, Temperature dependence of the CD signal from representative GB1 variants: ●, AAA-44Thr; +, AAA-44Val; ■, AAA-44Ala; ▲, AAA-44Asn; ◆, AAA-44Gly; and △, AAA-44Pro, in 150 mM NaCl, 50 mM Na-acetate, pH 5.4.

METHODS. As described previously 2 , major inter-residue contacts were identified 23 from the coordinate set 2GB1 and were defined as any atom pair having a centre-to-centre distance $\leq 150\%$ of the sum of the van der Waals radii. Major contacts to residue 44 arise from residue 53 on the adjacent strand with fewer, more distant contacts being made by residues at positions $i+2$ and $i-2$ from the guest site. These residues were changed to alanine to create the host molecule E42A/D46A/T53A (denoted AAA). To test for possible effects from the $i+2$ and $i-2$ residues, which had been serine in our central position study 2 , we also created the host molecule E42S/D46S/T53A (denoted SSA). Substitution at the guest position with threonine, an amino acid expected to be a very good β -sheet former, and glycine, an amino acid expected to be a very poor β -sheet former, yields identical $\Delta\Delta G$ values relative to alanine for β -sheet formation in both the AAA and SSA backgrounds. The host molecule bearing the smaller amino-acid substitutions, AAA, was chosen for further study. Recombinant GB1 mutants were expressed from a synthetic gene bearing the mutation Met 1 $\rightarrow$ Thr (GB1-Thr1) described previously 2 . Mutations were generated by single-strand mutagenesis 24 and were verified by dideoxyribonucleotide

sequencing (Sequenase, USB) of the entire mutant gene. Proteins were expressed in *Escherichia coli* (BL21 (DE3) pLysS) and were induced at an A_{600} of ~ 0.5 – 0.8 with a final concentration of isopropylthiogalactoside of 0.4 mM for 2–4 h. All folded proteins were purified by affinity chromatography 7 with IgG 6 Fast Flow Sepharose (Pharmacia) followed by reverse phase high pressure liquid chromatography (HPLC) purification on a Vydac preparative C18 column using a linear H $_2$ O-acetonitrile gradient in the presence of 0.1% TFA. The proteins are expressed as mixtures of N-terminally processed protein beginning with threonine at position 1 (56 residues) and non-processed protein (57 residues) beginning with methionine at position 0. The ratio of processed to unprocessed protein seems to depend on the stability of the molecule (data not shown). The HPLC purification step separates these two species and in all cases the 56-residue protein was used. For the unfolded molecule AAA-44Pro, the IgG affinity column step was replaced by G75 Sephadex chromatography in a buffer of 10 mM phosphate, 150 mM NaCl, pH 7.3. The identity of each HPLC purified protein was confirmed by laser desorption mass spectrometry (Finnigan Mat Laser-mat). All measured molecular weights were within 3 atomic mass units of the expected mass. CD measurements were made and thermal unfolding curves were fitted as described previously 2 . ΔC_p values for unfolding (data not shown) were similar ($\pm 15\%$) to those measured for the centre site AASS-Xaa molecules 2 . Protein concentration was determined by measuring absorbance of the unfolded protein 25 . Differential scanning calorimetry was performed with a Microcal MC-2 scanning calorimeter (Northampton, MA) as described previously 2 . AAA-44Thr, AAA-44Ser, AAA-44Gln, AAA-44Ala and AAA-44Gly were found to have $\Delta H_{van't Hoff}/\Delta H_{cal}$ ratios of 0.98, 0.94, 1.03, 0.98 and 0.99, respectively.



Our experiments indicate that β -sheet propensity is modulated strongly by tertiary context, even at solvent-exposed positions. This result provides an explanation for the differences in

[illegible]

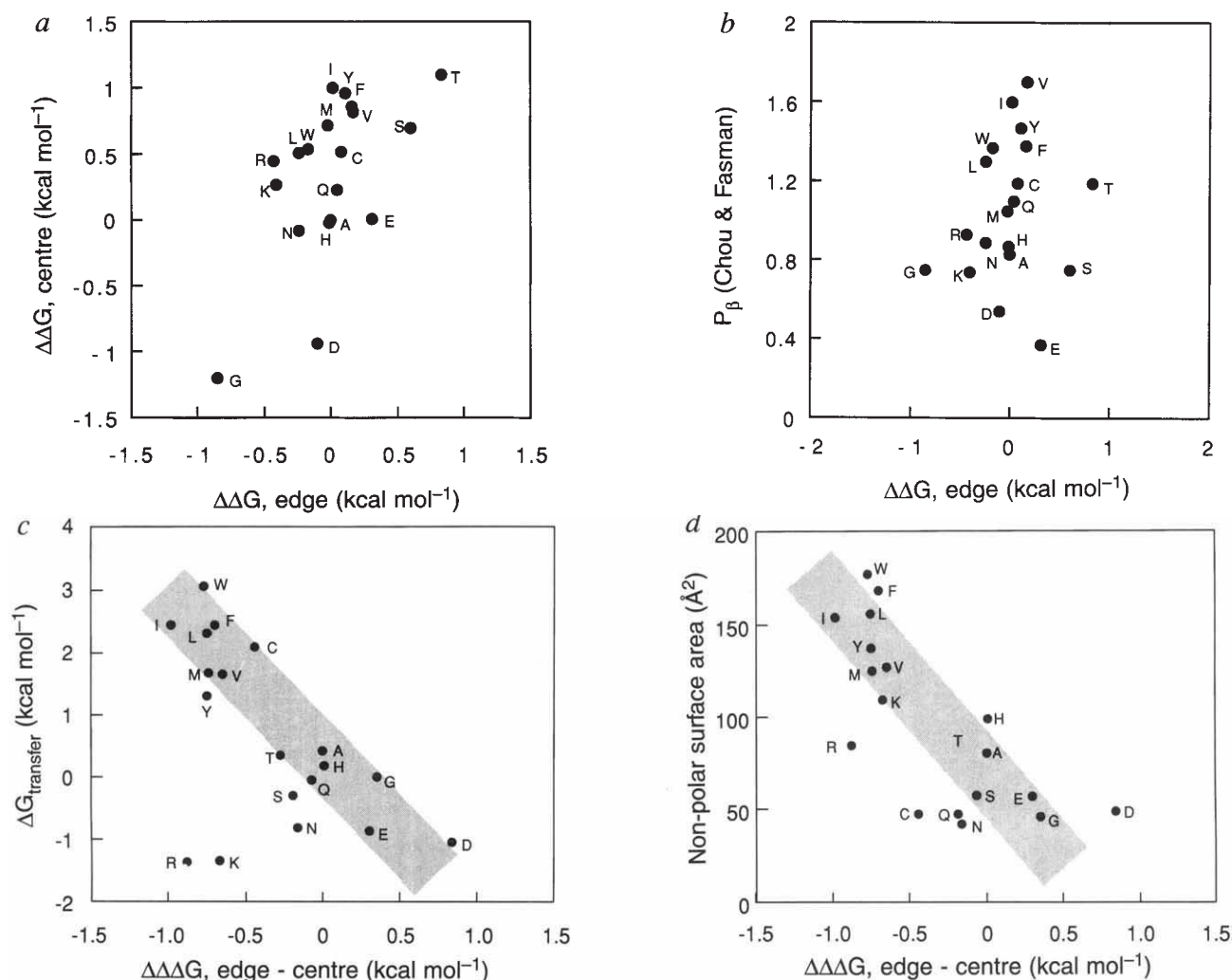


FIG. 3 *a*, Comparison of β -sheet forming propensities at central² and edge β -sheet positions. *b*, Comparison of the edge site β -sheet propensities with the β -sheet forming frequencies of Chou and Fasman⁴. The data are not correlated ($r=0.15$, $P>0.25$). *c*, Correlation of the difference in β -sheet forming propensity measured at edge and centre β -sheet positions, $\Delta\Delta\Delta G$, with ΔG of transfer for the amino acid side chains from octanol to water⁵ ($r=0.57$, $P<0.005$; $r=0.80$, $P<0.0005$, excluding values of R and K, see methods) and *d*, side chain non-polar surface area⁶ ($r=0.72$, $P<0.0005$). Amino acids are identified by the labels. Shaded areas are meant only to emphasize the overall trends in the data.

apparent β -sheet propensity measured in the zinc finger¹⁹ and GB1 (refs 2, 3) model systems. More generally, our results emphasize that β -sheets are elements of both secondary and tertiary structure²⁰. □

Received 13 June; accepted 2 August 1994.

- Chothia, C. *J. molec. Biol.* **105**, 1–14 (1976).
- Minor, D. L. Jr & Kim, P. S. *Nature* **367**, 660–663 (1994).
- Smith, C. K., Withka, J. M. & Regan, L. *Biochemistry* **33**, 5510–5517 (1994).
- Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211–222 (1973).
- Fauchere, J.-L. & Pliska, V. *Eur. J. med. Chem.-Chim. Ther.* **18**, 369–375 (1983).
- Livingstone, J. R., Spolar, R. S. & Record, M. T. Jr *Biochemistry* **30**, 4237–4244 (1991).
- Alexander, P., Fahnestock, S., Lee, T., Orban, J. & Bryan, P. *Biochemistry* **31**, 3597–3603 (1992).
- Frick, I.-M. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8532–8536 (1992).
- Gronenborn, A. M. & Clore, G. M. *J. molec. Biol.* **223**, 331–335 (1993).
- McGregor, M. J., Islam, S. A. & Sternberg, M. J. E. *J. molec. Biol.* **198**, 295–310 (1987).
- Lyu, P. C., Liff, M. I., Marky, L. A. & Kallenbach, N. R. *Science* **250**, 669–673 (1990).
- O'Neill, K. T. & DeGrado, W. F. *Science* **250**, 646–651 (1990).
- Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M. & Baldwin, R. L. *Nature* **344**, 268–270 (1990).
- Wojcik, J., Altman, K.-H. & Scheraga, H. A. *Biopolymers* **30**, 121–134 (1990).
- Chakrabarty, A., Schellman, J. A. & Baldwin, R. L. *Nature* **351**, 586–588 (1991).
- Horowitz, A., Matthews, J. M. & Fersht, A. R. *J. molec. Biol.* **227**, 560–568 (1992).
- Serrano, I., Neira, J.-L., Sancho, J. & Fersht, A. R. *Nature* **356**, 453–455 (1992).
- Blaber, M., Zhang, X. & Matthews, B. W. *Science* **260**, 1637–1640 (1993).
- Kim, C. A. & Berg, J. M. *Nature* **362**, 267–270 (1993).
- Lifson, S. & Sander, C. *J. molec. Biol.* **139**, 627–639 (1980).
- Priestle, J. P. *J. appl. Crystallogr.* **21**, 572–576 (1988).
- Gronenborn, A. M. et al. *Science* **253**, 657–661 (1991).
- Oas, T. G. & Kim, P. S. *Nature* **336**, 42–48 (1988).
- Kunkel, T. A., Roberts, J. D. & Zakour, R. A. *Meth. Enzym.* **154**, 367–382 (1987).
- Edelhoch, H. *Biochemistry* **6**, 1948–1954 (1967).
- Grathwohl, C. & Wüthrich, K. *Biopolymers* **20**, 2623–2633 (1981).
- McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G. & Dahlquist, F. W. *Biochemistry* **29**, 6341–6362 (1990).
- Wüthrich, K. *NMR of Proteins and Nucleic Acids* (John Wiley, New York, 1986).
- Rosner, B. *Fundamentals of Biostatistics* (Duxbury, Boston, 1982).
- Fahnestock, S. R., Alexander, P., Filipula, D. & Nagle, J. in *Bacterial Immunoglobulin-Binding Proteins* (ed. Boyle, M. D. P.) (Academic, San Diego, 1990).

METHODS. Values for amino acid side chain non-polar surface areas are for 'set 1' from ref. 14. The values from 'set 2' show a similar trend. With respect to the correlation between $\Delta\Delta\Delta G$ and ΔG of transfer there are two clear outlying points, arginine and lysine. The apparent anomalous behaviour of these two residues probably reflects the large distance between the side chain charge and the backbone β -sheet structure. Correlation coefficients were calculated for the (x_i, y_i) pairs using the formula²⁹:

$$r^2 = \left[\frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2 \right]^{1/2}} \right]^2$$

ACKNOWLEDGEMENTS. We thank Z.-Y. Peng for helpful advice and suggestions and L. M. Mayr, M. G. Oakley and P. A. Petillo critical reading of the manuscript. This work was supported by the Howard Hughes Medical Institute. D.L.M. is supported by a National Institutes of Health Biophysical Training Grant.

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2-3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50-80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963-65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.